US005113445A

United States Patent [19]

Wang

[11] Patent Number: 5,113,445

[45] Date of Patent: May 12, 1992

[54] SYSTEM FOR ENCODING DATA IN MACHINE READABLE GRAPHIC FORM

[75] Inventor: Yajiun P. Wang, Port Jefferson Station, N.Y.

[73] Assignee: Symbol Technologies Inc., Bohemia, N.Y.

[21] Appl. No.: 653,822

[22] Filed: Feb. 11, 1991

Related U.S. Application Data

[63] Continuation of Ser. No. 550,023, Jul. 9, 1990, abandoned.

[51] Int. Cl.[5] H04L 9/00

[52] U.S. Cl. 380/51; 235/462

[58] Field of Search 380/55, 51, 3, 59; 235/462

[56] References Cited

U.S. PATENT DOCUMENTS

3,647,275	3/1972	Ward	283/70
3,676,644	7/1972	Vaccaro et al.	235/449
3,761,683	9/1973	Rogers	235/380
3,812,328	5/1974	Tramposch	235/487
4,654,718	3/1987	Sueyoshi .	
4,654,793	3/1987	Elrod	235/380
4,660,221	4/1987	Dlugos	380/23
4,794,239	12/1988	Allais	235/462
4,835,713	5/1989	Pastor	380/51
4,863,196	9/1989	Ohnishi et al.	283/82
4,893,333	1/1990	Baran et al. .	
4,908,873	3/1990	Philibert et al. .	
4,949,381	8/1990	Pastor	380/51
4,975,948	12/1990	Andresen et al. .	

FOREIGN PATENT DOCUMENTS

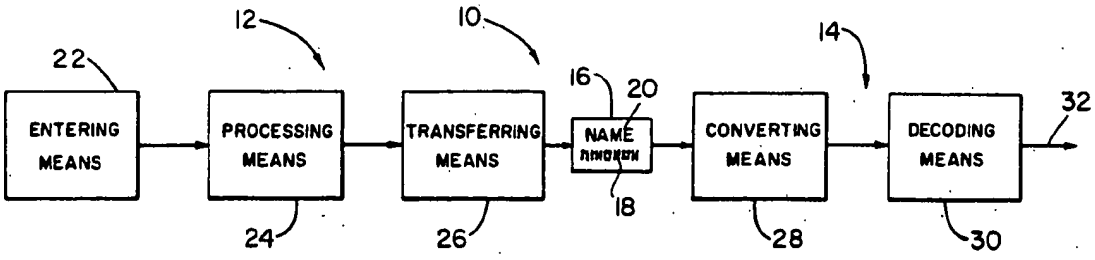
59-44157	3/1984	Japan .
62-91075	4/1991	Japan .

Primary Examiner—Thomas H. Tarcza
Assistant Examiner—David Cain

[57] ABSTRACT

A system for representing and recognizing data in machine readable graphic image form in which data to be encoded is entered into the system and a processor encodes the data into a two-dimensional pattern of graphic indicia and generates transfer drive signals representative of the indicia. A transferring device such as a printer transfers an image of the two-dimensional pattern of graphic indicia onto a carrier such as a card or paper document in responsde to the transfer drive signals. A recognition device converts the image on the carrier into electrical signals representative of the graphic indicia and decodes the signals into output signals representative of the data for further processing or use.

27 Claims, 3 Drawing Sheets



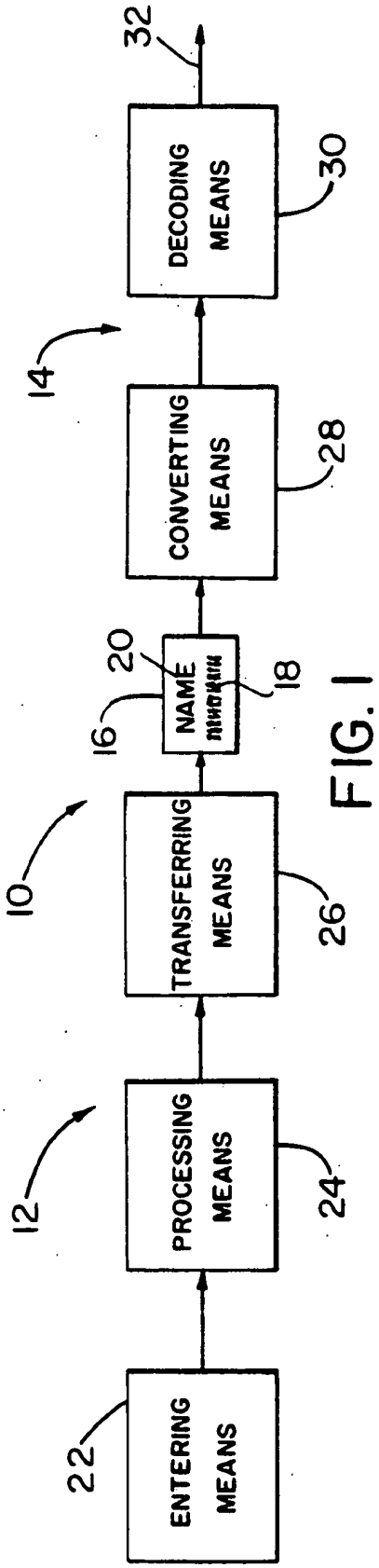

32
DECODING MEANS
30
14
CONVERTING MEANS
28
20
16
NAME
18
10
TRANSFERRING MEANS
26
12
PROCESSING MEANS
24
22
ENTERING MEANS

FIG. 1

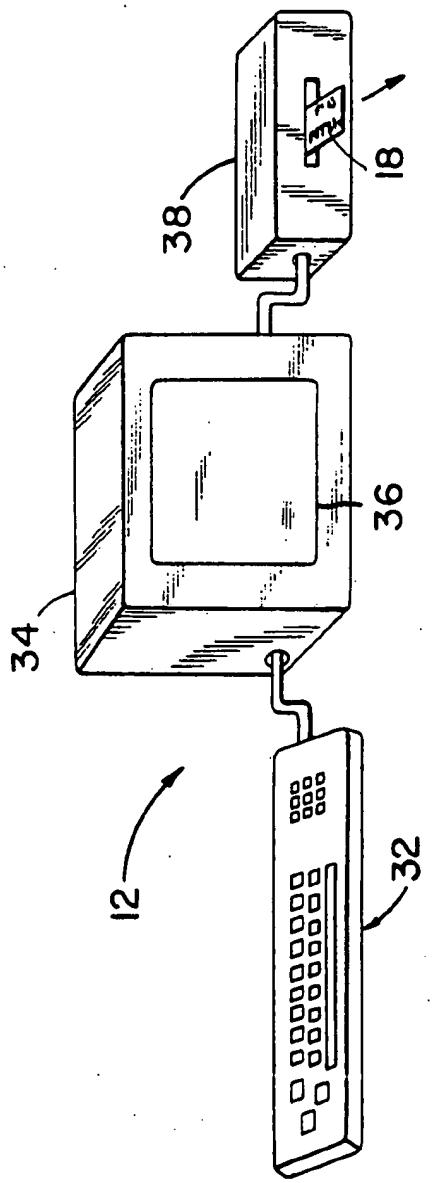

38
18
36
34
12
32

FIG. 2

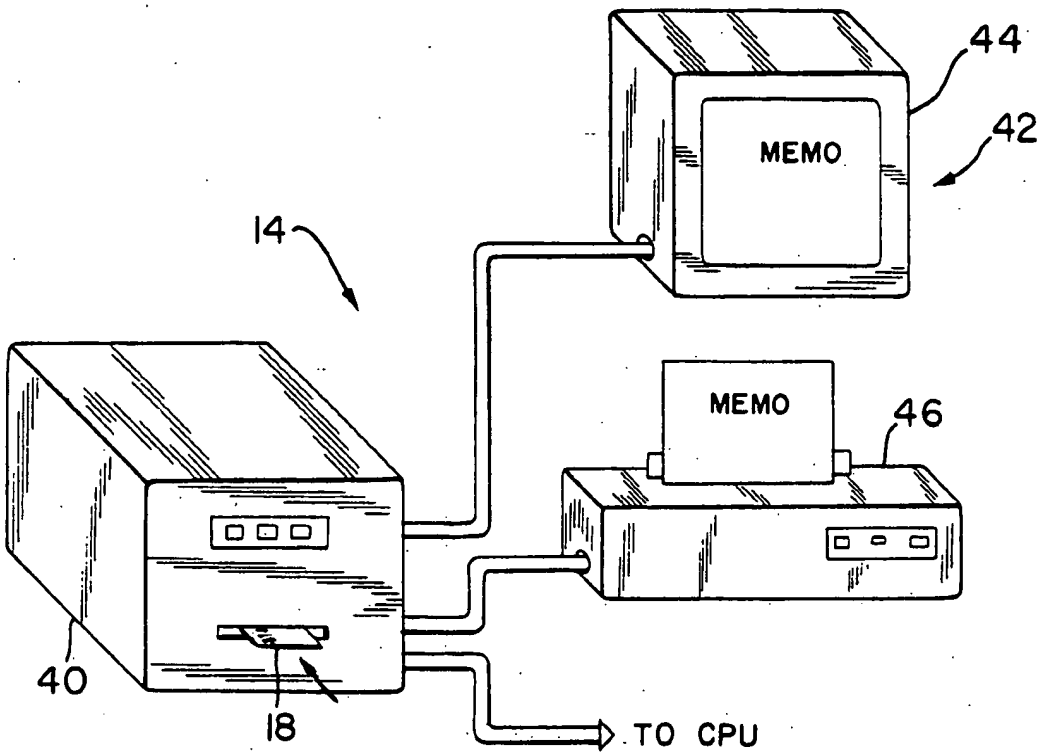


FIG. 3

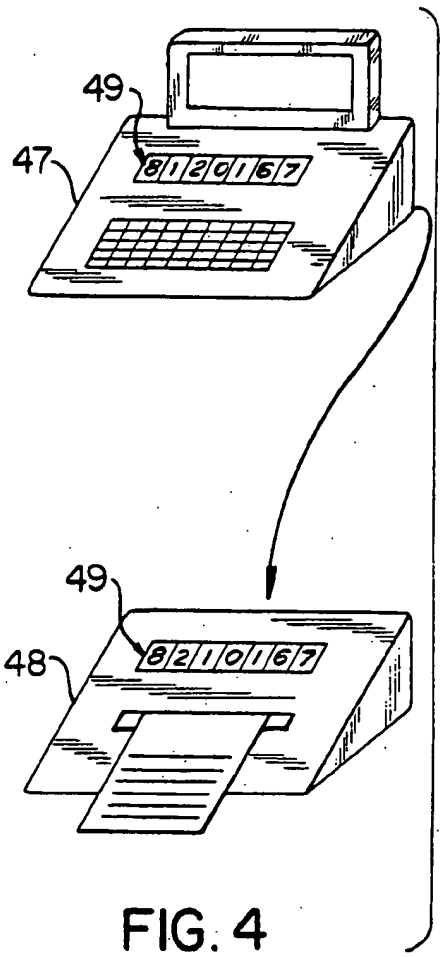


FIG. 4

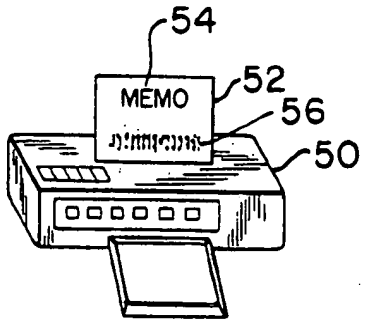
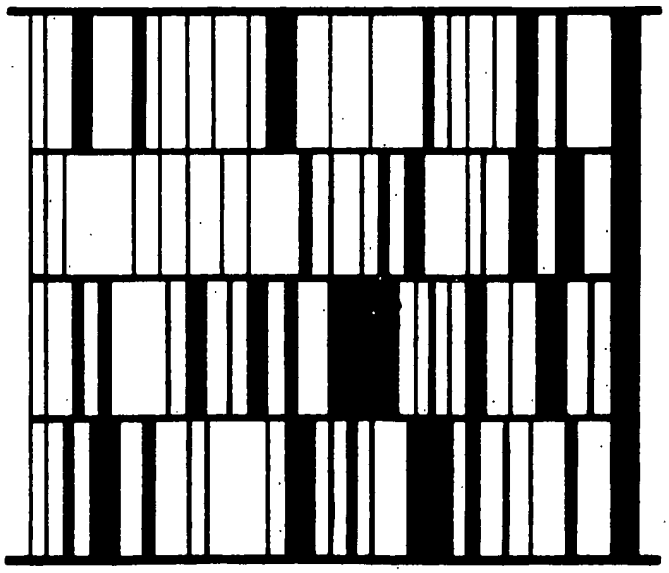


FIG. 5



PRIOR ART

FIG. 6

SYSTEM FOR ENCODING DATA IN MACHINE READABLE GRAPHIC FORM

This is a continuation of application Ser. No. 550,023 filed July 9, 1990 now abandoned.

REFERENCE TO RELATED APPLICATIONS

The present application is related to U.S. patent application Ser. No. 461,881 assigned to the same assignee as the present application.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention generally relates to the representation of data in machine readable form and more particularly to a system for the encryption of data into a two-dimensional graphic image that can be automatically machine read to obtain the encoded data in both open and closed systems.

2. Description of the Prior Art

In todays high technology world more and more operations are being automatically performed by machines and systems. This ever-increasing drive for automation has resulted in a demand for new techniques for encoding data into machine readable form for automatic entry into the various systems and machinery. The data entry may be for such uses as data transmission, operating various machine functions or the identification of persons or items. The various media that carry the data for automatic entry include punch cards, magnetic tapes and discs and magnetic stripes on cards such as Credit Cards and badges. The systems utilizing the above carriers are in "closed" systems, i.e. the read function is performed within an apparatus or housing and the reading element is in contact or in near-contact with the Carrier means during the reading operation.

One method for representing data in a machine readable form is to encode the data into a pattern of indicia having parts of different light reflectivity, for example, ar code symbols. A bar code symbol is a pattern comprised of a series of bars of various widths and spaced apart from one another by spaces of various widths, the bars and spaces having different light reflective properties. The bar code symbol is optically scanned and the resulting electrical signals are decoded into data representative of the symbol for further processing. Bar code reading systems are known as "open" systems in that the carrier while being read is not sealed, but is read from a distance and without being in physical contact with the scanner.

The conventional bar code described above is "one-dimensional" in that the information encoded therein is represented by the width of the bars and spaces, which extend in a single dimension. Similarly data encoded onto other media such as credit card magnetic stripes is composed of one or more "one-dimensional" tracks of encoded data.

The use of bar code symbols and magnetically encoded data has found wide acceptance in almost every type of industry. However, the one dimensional nature of the encoded data limits the amount of information that can be encoded and hence use has been generally restricted to simple digital representations.

Thus, there is a need for a system to encode data in machine readable form that allows for an increase in the amount of data encoded into a given space that can be quickly and easily decoded for further processing.

SUMMARY OF THE INVENTION

The present invention is directed to a system for representing and recognizing data in machine readable graphic image form having an increased capacity for encoded information that can be used in both open and closed systems. The system comprises an encoding means having a means for entering data such as a keyboard or optical character scanner. In addition, the data may be obtained directly from computer files. The data entered into the system may be both textual data and control data. The data is entered into a processing means for encoding the data into a two-dimensional pattern of graphic indicia. The graphic indicia may, for example, be in the form of a two-dimensional bar code which is comprised of a pattern of vertical bars of predetermined lengths that are spaced at various vertical and horizontal intervals. It should be understood that the graphic indicia representative of the data is not limited to a bar code symbol type, but may be in the form of any two-dimensional graphic pattern of indicia suitable for encoding data.

The processing means generates electrical drive signals for transferring the two-dimensional graphic pattern onto a data Carrier means, that may be a card or document or the surface of a machine part. The encoding means also includes means for transferring an image of the two-dimensional pattern of graphic indicia onto the data carrier means in response to the transfer drive signals.

The image may for example be printed in the form of a two-dimensional pattern of graphic indicia having different areas of light reflectivity in which the indicia have one level of reflectivity and the spaces have another level of reflectivity. In this embodiment, the converting means may be a type of optical scanner typically used for scanning one-dimensional bar codes that Converts the areas of different light reflectivity into electrical signals representative of the indicia. Scanners employed in the present invention, however, have the added feature of scanning the indicia in two dimensions. For example, in one method a laser light beam is scanner across the indicia in a raster pattern for reading and decoding two-dimensional graphic codes. Optical scanners suitable for reading two-dimensional patterns are disclosed in U.S. patent application Ser. Nos. 317,433 and 317,533, filed Mar. 1, 1989, assigned to the same assignee as the present invention and incorporated herein by reference.

The system of the present invention further includes a recognition means comprising means for converting the image on the carrier means into electrical signals representative of the graphic indicia and means for decoding the electrical signals into output signals representative of the data.

The decoded output signals are available for further processing and the system may therefore include means for outputting the decoder output signals. Typical output devices may include a liquid crystal display, a CRT display and a printer. The outputted signals may also be transmitted to a computer or other system for further processing and use via telephone lines using a modem or via a data bus.

The present invention contemplates the outputting of the decoder output signals to a microprocessor for controlling the operation of various machines such as facsimile, VCR, microwave oven, robotic systems and weight/price label scale devices.

In another embodiment of the invention, the processing means encodes a first set of data into the two-dimensional pattern of graphic indicia and generates first transfer drive signals for transferring the two-dimensional pattern onto a carrier means. The processing means also generates a second set of transfer drive signals in response to a second set of data entered into the system intended to be transferred to the carrier means in human readable form. Thereafter, the transfer means transfers onto the carrier means both the image of the two-dimensional graphic pattern of indicia in response to the first transfer drive signals and the second set of data in human readable form in response to the second transfer drive signals. Thus, the system provides means for automatically representing data in both a machine readable form and human readable form onto a single carrier means.

In yet another embodiment of the invention, the data is encoded and decoded using a keyed data encryption technique in order to increase the security of the data transmission. In this embodiment, only the person having the encryption key will be able to decode the graphic pattern.

The system of the present invention maximizes the use of available space for encrypting data. In addition to being compact in size, the system provides for high security in the transmission of information. Thus, the invention provides a highly reliable system for representating data in machine readable graphic form having increased encoding capacity thereby substantially expanding applications for automatic data entry. In addition, the invention creates a new media for man-machine interfacing.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of the system of the present invention.

FIG. 2 is a perspective view of an encoding means of the system of the present invention.

FIG. 3 is a perspective view of a recognition means of the system of the present invention.

FIG. 4 is a perspective view of a data entry device and a reader in which a key may be entered for encrypting and decrypting data.

FIG. 5 is a perspective view of a facsimile machine incorporating the recognition means of the present invention.

FIG. 6 is a representation of an embodiment of a two dimensional bar code as the term is used in this patent specification.

DETAILED DESCRIPTION OF THE INVENTION

Referring now to the drawings, FIG. 1 is a block diagram of the system 10 of the present invention for representing and recognizing data in machine readable graphic image form. System 10 includes an encoding means generally indicated by the reference numeral 12 and a recognition means generally indicated by the reference numeral 14. Encoding means 12 produces a carrier means 16 containing at least a two-dimensional pattern of graphic indicia 18. Carrier means 16 may also contain human readable data 20. The two-dimensional pattern of graphic indicia on carrier means 16 is recognized by recognition means 14 to produce output signals representative of the data encoded into the pattern 18.

Data to be transferred onto carrier means 16 is entered by entering means 22 into the encoding means 12. The data entered by entering means 22 may be both the data to be encoded into the two-dimensional pattern of graphic indicia and the data to appear on carrier means 16 in human readable form. Processing means 24 encodes the set of data to appear in pattern 18 into a two-dimensional pattern of graphic indicia and generates transfer drive signals for controlling the transfer of the indicia onto the carrier means 16. Transferring means 26 transfers an image of the two-dimensional pattern of graphic indicia onto carrier means 16 in response to the transfer drive signals. If human readable data is also to be transferred onto carrier 16, the processing means 24 generates a second set of transfer drive signals for controlling the transfer of the human readable data onto carrier 16. A portion or all of the data to be encoded and the human readable data may be transferred from a storage memory in processing means 24 or other computer files rather than being entered by means 22.

The carrier means 16 is shown in FIGS. 1, 2 and 3 is represented as being in the form of a card approximately the size of a credit card. This type of card is illustrative only as the carrier means 18 may be made of any material on which graphic indicia may be transferred to, such as paper, etc.

Recognition means 14 includes converting means 28 that converts the image on carrier means 16 into electrical signals representative of the graphic indicia. Decoding means 30 decodes the electrical signals into decoder output signals indicated at 32 that are representative of the data encoded into the pattern 18.

FIG. 2 is a perspective view of one embodiment of encoding means 12. In this embodiment, the entering means 22 of FIG. 1 is shown in form of a keyboard 32 for entering alphanumeric and graphic data into the decoding means 12. The embodiment of FIG. 2 is illustrative only as entering means 22 may take forms other than a keyboard such as an optical scanning means for scanning data directly from documents for entry into the encoding means 12. Entering means 22 may also be in the form of various card readers in which magnetically encoded information is scanned and converted into electrical signals representative of the data.

Referring again to FIG. 2, the processing means 24 of FIG. 1 is shown in the form of a processor and display unit 34. The data entered by keyboard 32 is transmitted to the processor and display unit 34 for storage and processing. In addition to entering data, the keyboard 32 is also used for entering control commands to effect operation of the processor unit 34.

The data entered by keyboard 32 is displayed on display screen 36 and upon entry of a proper control command, is also stored in memory. The data to be encoded into the pattern of graphic indicia is stored in a first memory, in processor 34 and the data, if any, to be transferred in human readable form is stored in a second memory. Alternatively, the data may be stored in a separate portion of a single memory. Upon the appropriate control command from keyboard 32, the processor unit 34 encodes the data in the first memory into a two-dimensional pattern of graphic indicia and generates first transfer drive signals representative of the data stored in the first memory. The processor unit 34 also generates second transfer drive signals representative of the data stored in the second memory.

The processor unit 34 is shown in FIG. 2 as being coupled to a printer 38. The printer 38 is one form of the transferring means 26 of FIG. 1. Printer 38 transfers an image of the two-dimensional pattern of graphic indicia

on carrier means 18 in response to the first transfer drive signals and prints the second set of data in human readable form onto carrier means 18 in response to the second transfer drive signals. In one embodiment, the printer 38 prints the two-dimensional pattern in the form of graphic indicia having different areas of light reflectivity, such as a two-dimensional bar code. Printer 38 may take other forms such as a means for printing the two-dimensional pattern of graphic indicia with magnetic-ink. In such a device, magnetic indicia are deposited on the carrier material in a two-dimensional pattern that may be recognized by magnetic-ink recognition sensors.

Turning now to FIG. 3, the recognition means 14 includes a card reader 40 which contains the converting means 28 and the decoding means 30 of FIG. 1. The converting means 28 may be a bar code reader such as those disclosed in U.S. patent application Ser. Nos. 317,433 and 317,533, assigned to the same assignee as the present invention and incorporated herein by reference. The readers disclosed in the above patent applications are open system devices designed to read an optically encoded two-dimensional bar code and to convert the light reflected from the pattern into electrical signals representative of the graphic indicia.

The card reader 40 may also comprise a magnetic-ink recognition device for reading and decoding magnetically encoded data. These closed system devices include a magnetic read head that senses the change in reluctance associated with the presence of the magnetic-ink. The use of appropriate converting means that corresponds to the particular data encoding technology employed is contemplated by the present invention.

The decoding means 30 decodes the electrical signals into output signals representative of the data encoded onto carrier means 18. The decoder output signals are outputted from the recognition unit 40 to various output means 42. FIG. 3 depicts two examples of output devices; one being a display unit 44 and the other a printer 46. Display unit 44 may be any suitable display such as liquid crystal display or a CRT. The printer 46 may be any print device such as a dot matrix printer, laser printer, etc.

The system of the present invention maximizes the use of available space for encrypting data. The density of the encoded data is such that for a two-dimensional bar code symbol, a minimum of about 1600 characters can be encoded into a space of approximately 5"×½". In addition to being compact in size, the system provides for high security in the transmission of information. For example, a sensitive message may be encoded onto a document also containing non-sensitive material. This document, the same as any document, can be copied, transmitted by facsimile, etc., but only those with a recognition means of the present invention will be able to "read" the sensitive portion. The carrier means, being a single sheet of paper or a plastic credit card type of card, is an inexpensive read-only-memory structure that facilitates data communication.

In another embodiment, the data may be encoded using a keyed encryption algorithm that may be accessed only by an encryption key. As shown in FIG. 4, the data entry means 47 contains the keyed algorithm and upon entry of the key 49, the data will be encoded into a two-dimensional graphic pattern in a unique configuration. The unique configuration can only be read by a reader 48 having the algorithm and only upon entry of the key 49 into the reader. Thus, a high degree of security may be provided with the keyed encryption embodiment.

In addition, the recognition unit 40 may also transmit the output signals to a central processing unit locally or remotely, by for example a modem, for further use or processing by the CPU. In this embodiment, the data encoded onto the carrier means 18 may be control data in the form of machine operating instructions for controlling a robotic system or to a security identification system for performing such functions as unlocking doors. In connection with the use of the present invention in a robotic system, it is contemplated that the two-dimensional graphic pattern containing the Control data be placed or printed directly onto a machine part or part holder. A scanner coupled to the machine tool reads the pattern and transmits the decoded instruction to the control computer which in turn controls the machining of the part in accordance with the control program. Another example of the use of the present invention includes a microwave food container where the two-dimensional graphic pattern contains instructions automatically entering the recommended cooking sequence. A further use may be in connection with placing on roadway signs two-dimensional patterns containing geographic location information that may be read by a scanner in passing vehicles for use with onboard computers.

The present invention further contemplates the use of the system of the present invention to encode control data containing machine operating instructions onto the carrier means in the form of machine readable graphic indicia that may be inserted into the machine to effect operation of the machine. FIG. 5 is an example of a facsimile machine 50 in which a document 52 containing human readable data 54 and a two-dimensional pattern of graphic machine readable indicia 56. The document 52 is inserted into the facsimile machine 50 the same as documents are normally inserted for transmission. The machine 50 contains a converting means for converting the two-dimensional image into electrical signals and a decoding means for decoding the signals into output signals operative to actuate the facsimile machine 50. The pattern 56 may contain such information as the phone number of the intended recipient of the memo 54 and the appropriate instructions for automatically entering the phone number and actuating the transmission process. Thus, where numerous messages are faxed to a particular recipient a supply of paper containing the phone number of the recipient encoded in the two-dimensional graphic indicia machine readable format may be maintained by the sender. The transmission of messages to that recipient will be facilitated by placing the message onto the pre-encoded paper and simply inserting the paper into the facsimile machine. In addition to simplifying and speeding the transmission process, the possibility of sending highly sensitive information to an incorrect party will also be eliminated.

FIG. 6 is a representation of an embodiment of a two dimensional bar code as the term is used in this patent specification. Such a bar code, taken from Allais, U.S. Pat. No. 4,794,239, includes a plurality of ordered rows of codewords of bar code information, one row adjacent to and beneath another row, each codeword representing at least one information bearing character.

While the invention has been particularly shown and described with respect to preferred embodiments thereof, it will be understood by those skilled in the art that the foregoing and other changes in form and detail

may be made therein without departing from the spirit and scope of the invention which should be limited only by the scope of the appended claims.

What is claimed is:

1. A system for representing and recognizing data on a record carrier in the form of a machine readable two-dimensional bar code structure comprising:

(a) encoding means including:

(i) means for entering data in said encoding means;

(ii) processing means for encoding said data into a two-dimensional bar code structure, said bar code structure including a plurality of ordered, adjacent rows of codewords of bar-coded information, each of said codewords representing at least one information-bearing character, and for generating first transfer drive signals;

(iii) means for transferring an image of the two-dimensional bar code structure onto a portable record carrier in response to said first transfer drive signals; and

(b) recognition means including:

(i) means for scanning the image of the two-dimensional bar code structure and for converting the codewords into electrical signals representative of the information-bearing characters; and

(ii) means for decoding the electrical signals into output signals representative of said data.

2. The system of claim 1 wherein said recognition means includes an output device for displaying said data in human readable form.

3. The system of claim 2 wherein said output device is a liquid crystal display.

4. The system of claim 2 wherein said output device is a CRT display.

5. The system of claim 2 wherein said output device is a printer.

6. The system of claim 1 wherein said recognition means includes means for transmitting the decoder output signals to a computer.

7. The system of claim 6 wherein said transmitting means includes a modem.

8. The system of claim 1 wherein said recognition means includes means for transmitting the decoder output signals to a microprocessor.

9. The system of claim 8 wherein the microprocessor controls the operation of one of a facsimile machine, a VCR, a microwave oven, a robot and a weight/price label scale, in response to said decoder output signals.

10. The system of claim 11 wherein said entering means includes a keyboard for entering said data.

11. The system of claim 1 wherein said entering means includes means for scanning data from a sheet.

12. The system of claim 11 wherein said processing means includes means for generating second transfer drive signals in response to a second set of data, and wherein said transferring means includes reading means for transferring onto said carrier means both the image of the two-dimensional bar code structure in response to the first transfer drive signals and said second set of data in human readable form in response to said second transfer drive signals.

13. The system of claim 1 wherein the transferring means is a printer.

14. The system of claim 12 wherein said portable record carrier is a single carrier, and wherein said reading means includes signal means for transferring the two-dimensional bar code structure and the second set of data in human readable form onto the portable carrier.

15. A method for representing and recognizing data on a record carrier in the form of a machine readable two-dimensional bar code structure comprising the steps of:

(i) entering said data into an encoding station;

(ii) encoding said data into a two-dimensional bar code structure, said bar code structure including a plurality of ordered, adjacent rows of codewords of bar-coded information, each of said codewords representing at least one information-bearing character;

(iii) transferring an image of the two-dimensional bar code structure onto a portable record carrier;

(iv) scanning the image of the two-dimensional bar code structure in a separate decoding station and converting the codewords into electrical signals representative of the information-bearing characters; and

(v) decoding the electrical signals into output signals representing said data.

16. The method of claim 15 wherein said entering step includes the substep of entering said data on a keyboard.

17. The method of claim 15 wherein said entering step includes the substep of scanning said data from a sheet.

18. The method of claim 15 wherein said transferring step includes the substep of printing the image onto the carrier.

19. The method of claim 15 further comprising the step of displaying the data in human readable form.

20. The method of claim 19 wherein said displaying step includes the substep of displaying the data on a CRT display.

21. The method of claim 19 wherein said displaying step includes the substep of displaying the data on a liquid crystal display.

22. The method of claim 19 wherein said displaying step includes the substep of printing the data on a printer.

23. The method of claim 15 further comprising the step of transmitting the output signals to a computer.

24. The method of claim 23 wherein said transmitting step includes the substep of transmitting the output signals to the computer by a modem.

25. The method of claim 15 further comprising the step of transmitting the output signals to a microprocessor.

26. The method of claim 25 further comprising the step of controlling the operation of one of a facsimile machine, a VCR, a microwave oven, a robot and a weight/price label scale, in response to the output signals.

27. The method of claim 15 further comprising the steps of

entering a second set of data into the encoding station; and

transferring onto the portable record carrier both the image of the two-dimensional bar code structure and said second set of data in human readable form.

* * * * *

US005337362A

United States Patent [19]

Gormish et al.

[11] Patent Number: 5,337,362

[45] Date of Patent: Aug. 9, 1994

[54] METHOD AND APPARATUS FOR PLACING DATA ONTO PLAIN PAPER

[75] Inventors: Michael J. Gormish, Stanford; Mark Peairs, Menlo Park; David G. Stork, Stanford, all of Calif.

[73] Assignees: Ricoh Corporation, Menlo Park, Calif.; Ricoh Company Ltd., Tokyo, Japan

[21] Appl. No.: 48,376

[22] Filed: Apr. 15, 1993

[51] Int. Cl.[5] G09C 5/00

[52] U.S. Cl. 380/54; 380/30

[58] Field of Search 380/29, 30, 46, 51, 380/54, 55, 59; 235/494

[56] References Cited

U.S. PATENT DOCUMENTS

4,782,221	11/1988	Brass et al. .	
4,853,961	8/1989	Pastor	380/30 X
4,939,354	7/1990	Priddy et al.	235/496 X
5,113,445	5/1992	Wang .	
5,170,044	12/1992	Pastor	235/496 X

FOREIGN PATENT DOCUMENTS

0469864	7/1991	European Pat. Off. .
0493053	12/1991	European Pat. Off. .
0493091	12/1991	European Pat. Off. .

Primary Examiner—Tod R. Swann

Attorney, Agent, or Firm—Blakely, Sokoloff, Taylor & Zafman

[57] ABSTRACT

A method and apparatus for placing digital data on plain paper. One embodiment of the present invention allows for the digital data to undergo encryption before being placed on the plain paper. In one embodiment, a photocopier is used for transferring digital encrypted data to and from a plain piece of paper. The photocopier allows digital data to be stored onto plain paper after encryption, such that the digital data is secure. The photocopier also includes a device to recognize the encrypted digitized pixels on the page such that they may be decrypted and the original image reproduced.

37 Claims, 9 Drawing Sheets

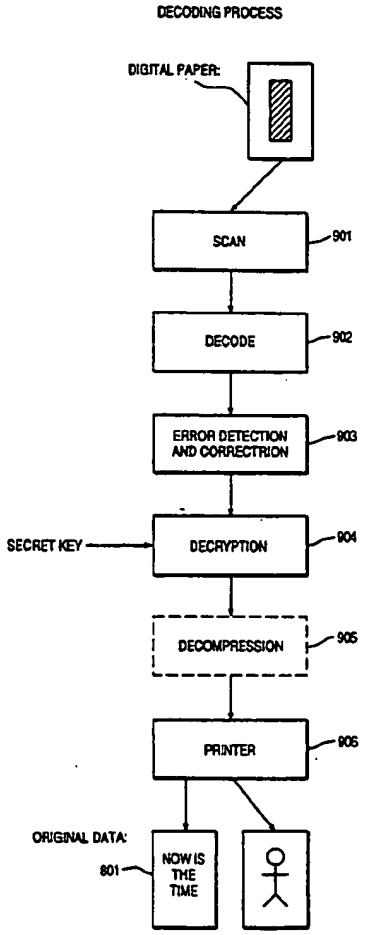
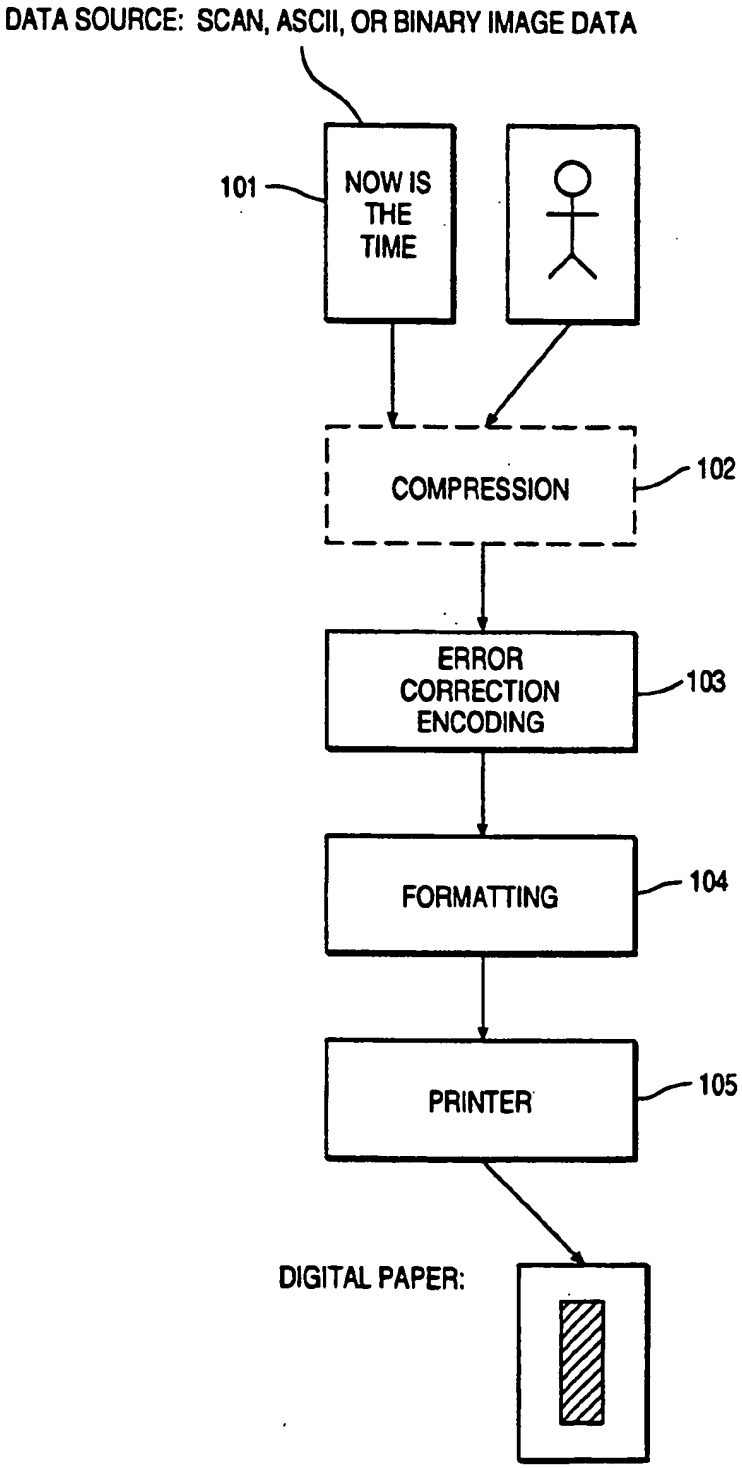


FIG_1 ENCODING PROCESS



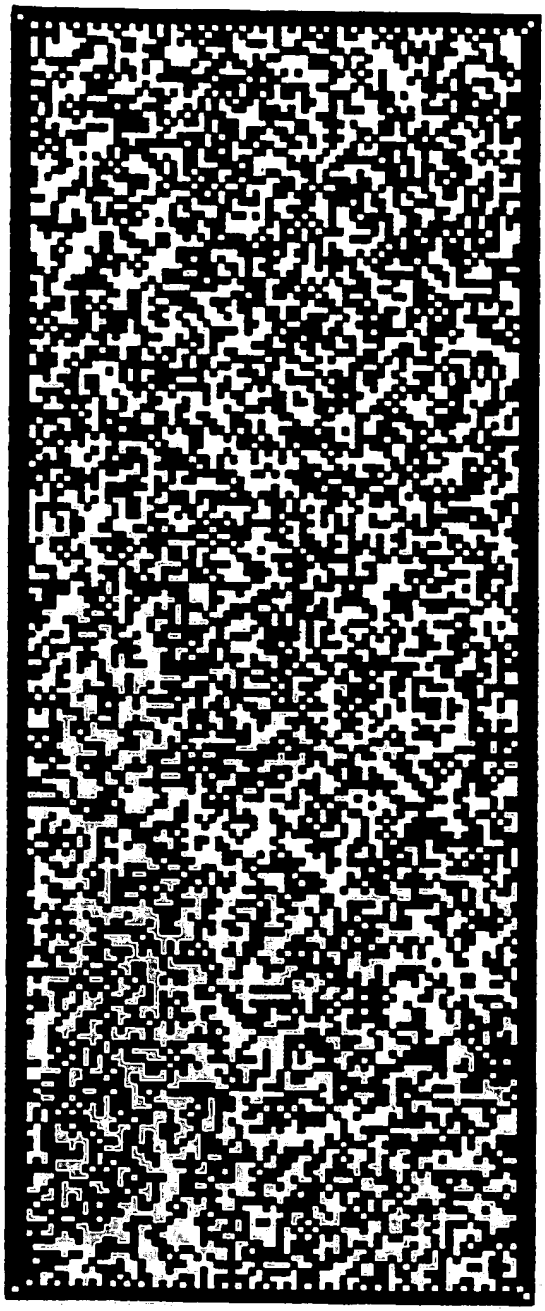


FIG_2

FIG_3

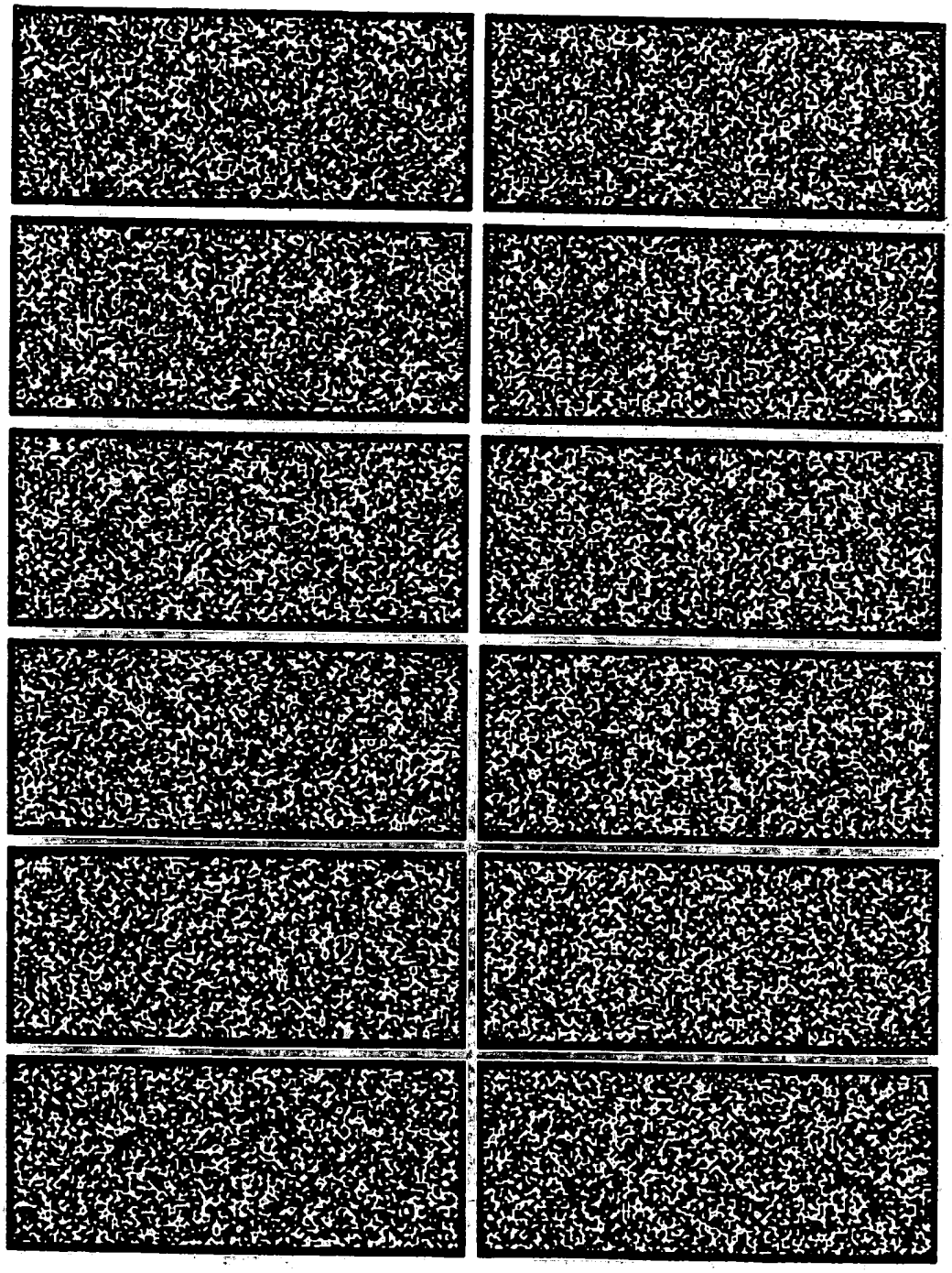
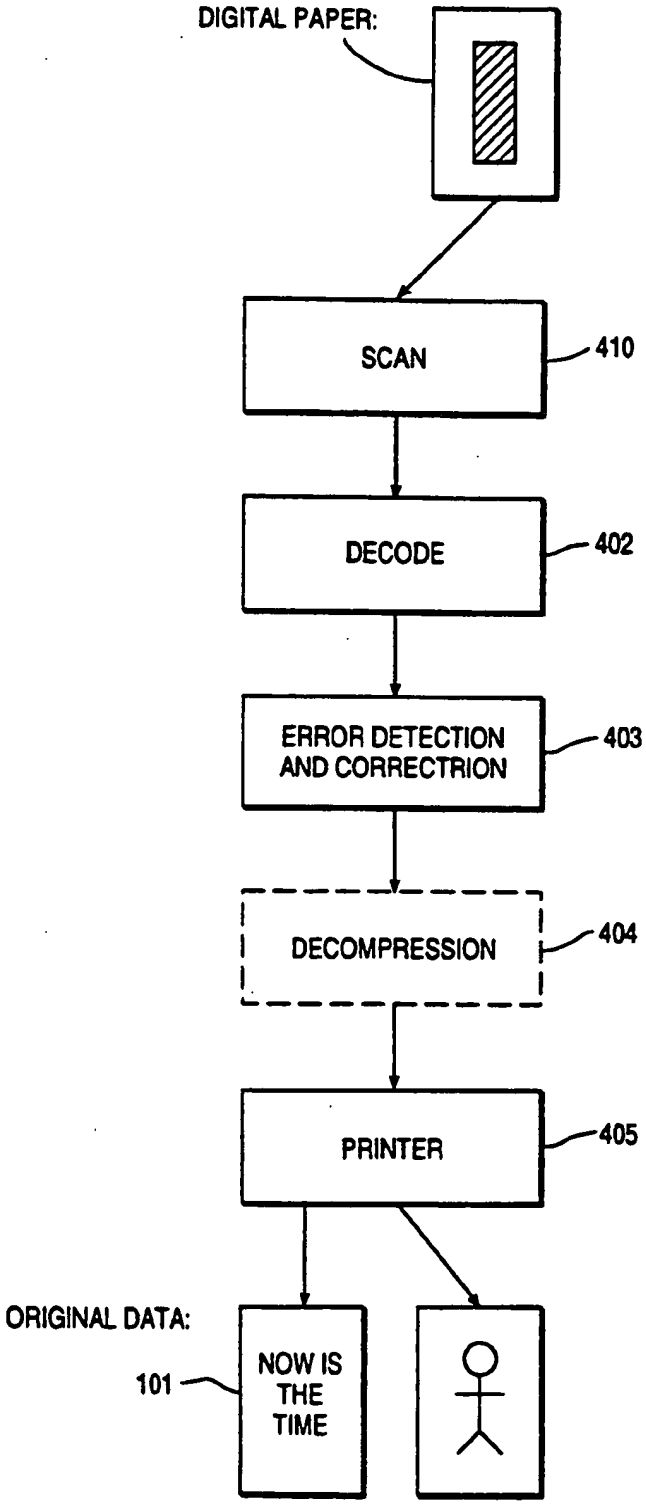


FIG_4 DECODING PROCESS



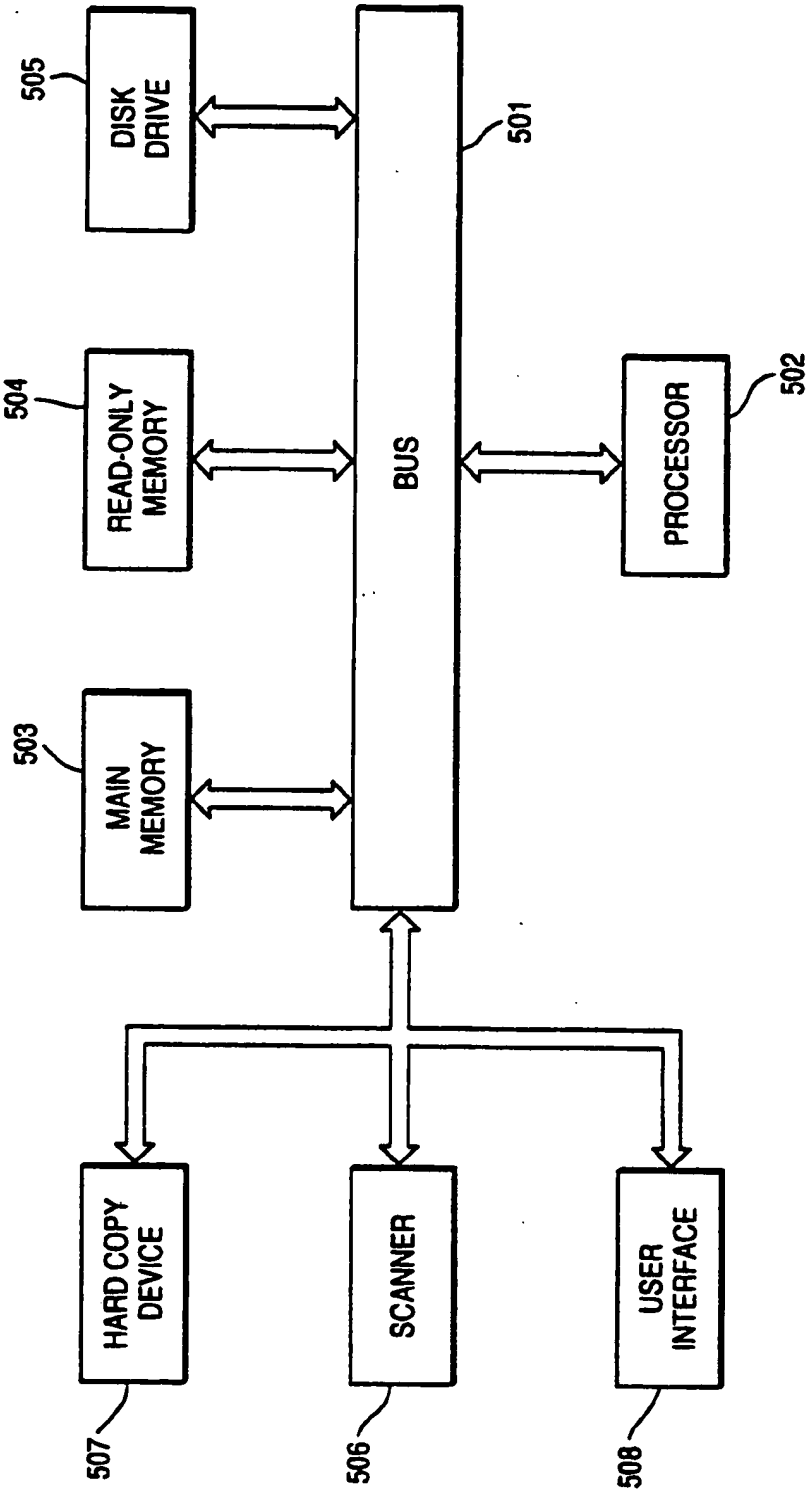
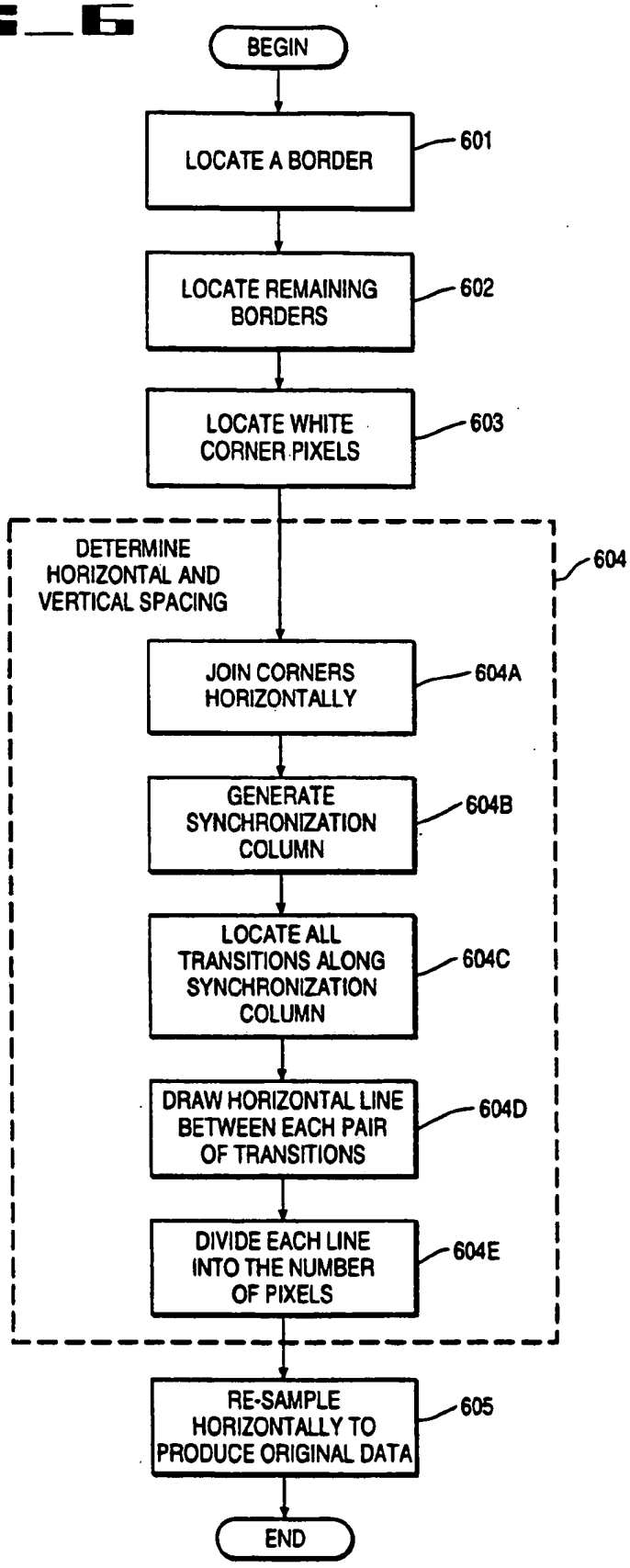

505
DISK DRIVE
501
504
READ-ONLY MEMORY
BUS
PROCESSOR
502
503
MAIN MEMORY
HARD COPY DEVICE
SCANNER
USER INTERFACE
507
506
508

FIG_5

FIG_6


BEGIN
LOCATE A BORDER
601
LOCATE REMAINING BORDERS
602
LOCATE WHITE CORNER PIXELS
603
DETERMINE HORIZONTAL AND VERTICAL SPACING
604
JOIN CORNERS HORIZONTALLY
604A
GENERATE SYNCHRONIZATION COLUMN
604B
LOCATE ALL TRANSITIONS ALONG SYNCHRONIZATION COLUMN
604C
DRAW HORIZONTAL LINE BETWEEN EACH PAIR OF TRANSITIONS
604D
DIVIDE EACH LINE INTO THE NUMBER OF PIXELS
604E
RE-SAMPLE HORIZONTALLY TO PRODUCE ORIGINAL DATA
605
END

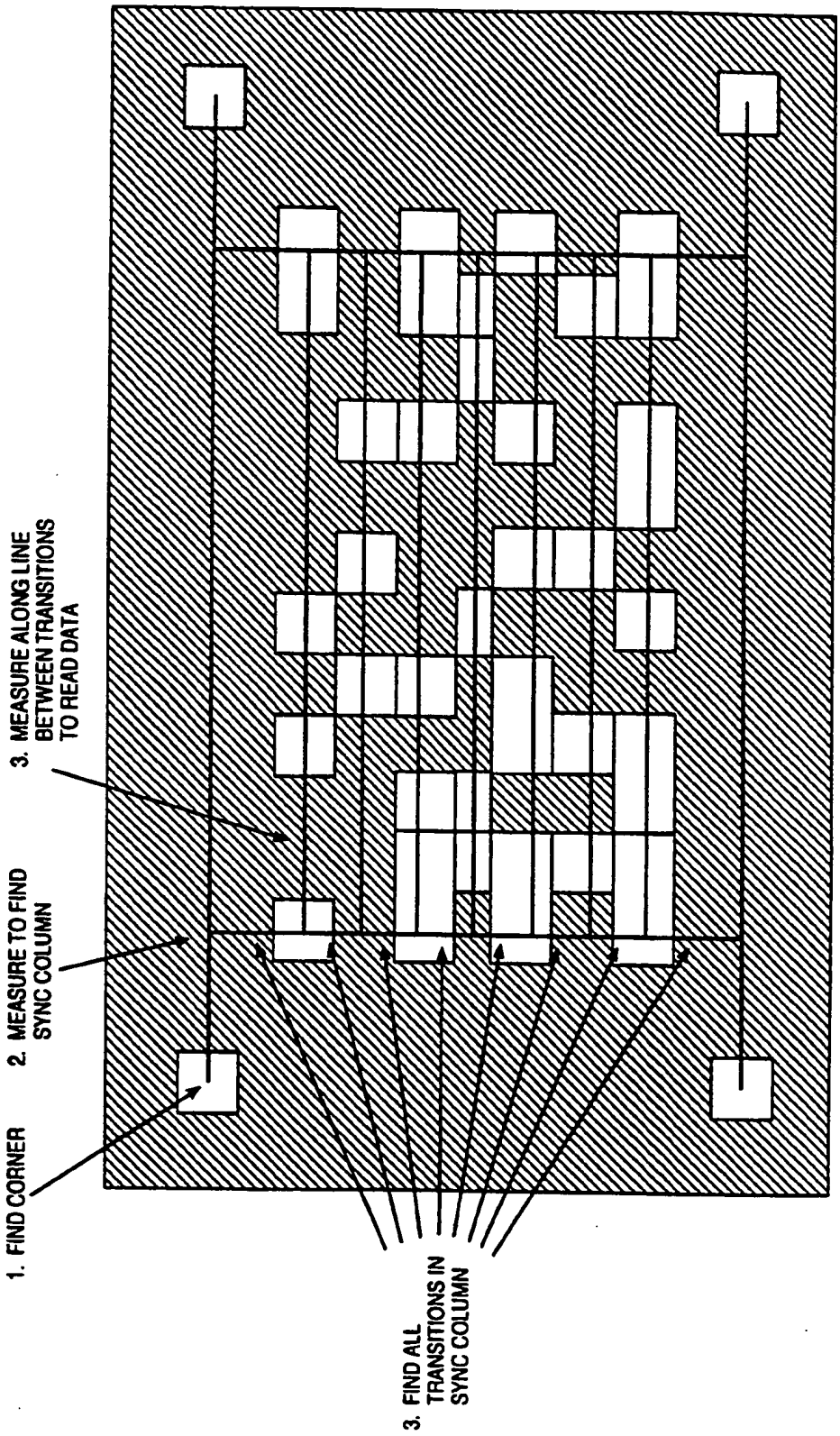

1. FIND CORNER
2. MEASURE TO FIND SYNC COLUMN
3. MEASURE ALONG LINE BETWEEN TRANSITIONS TO READ DATA
3. FIND ALL TRANSITIONS IN SYNC COLUMN

FIG_7

FIG_8 ENCODING PROCESS

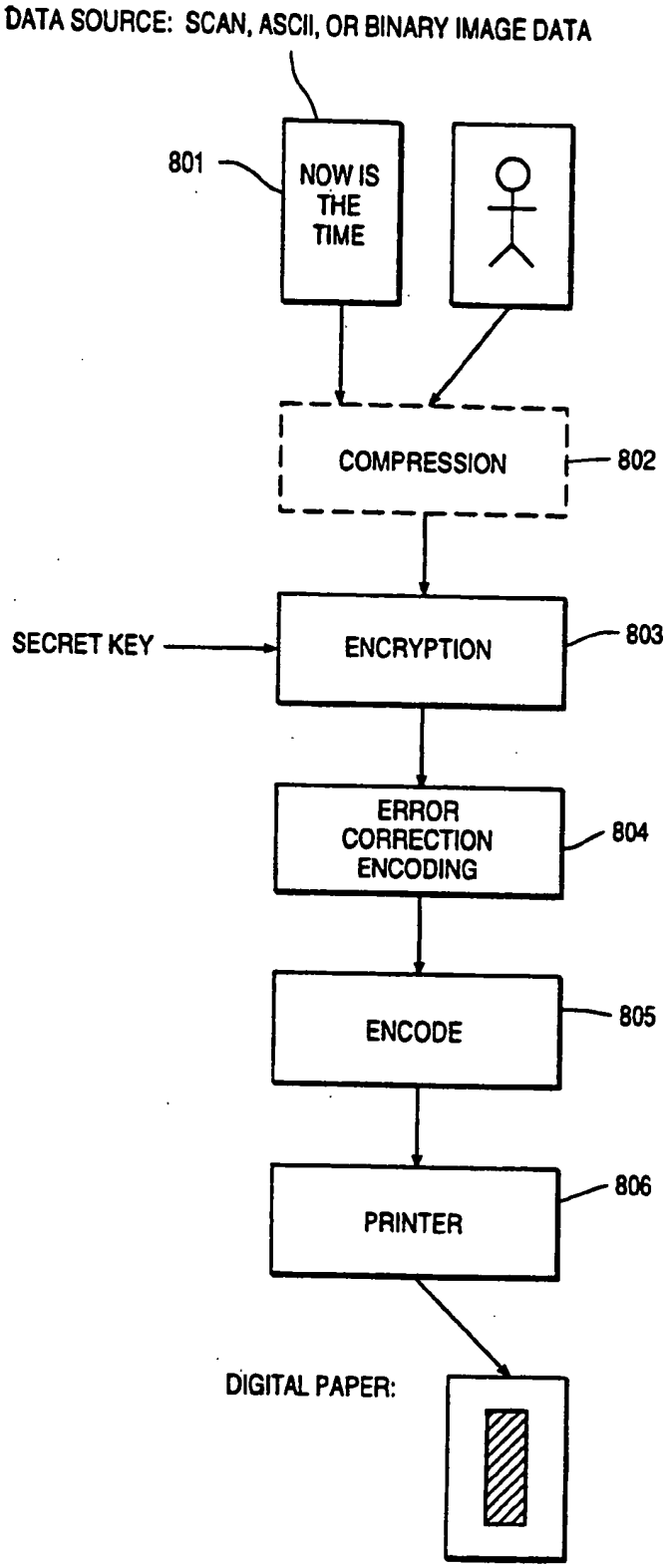
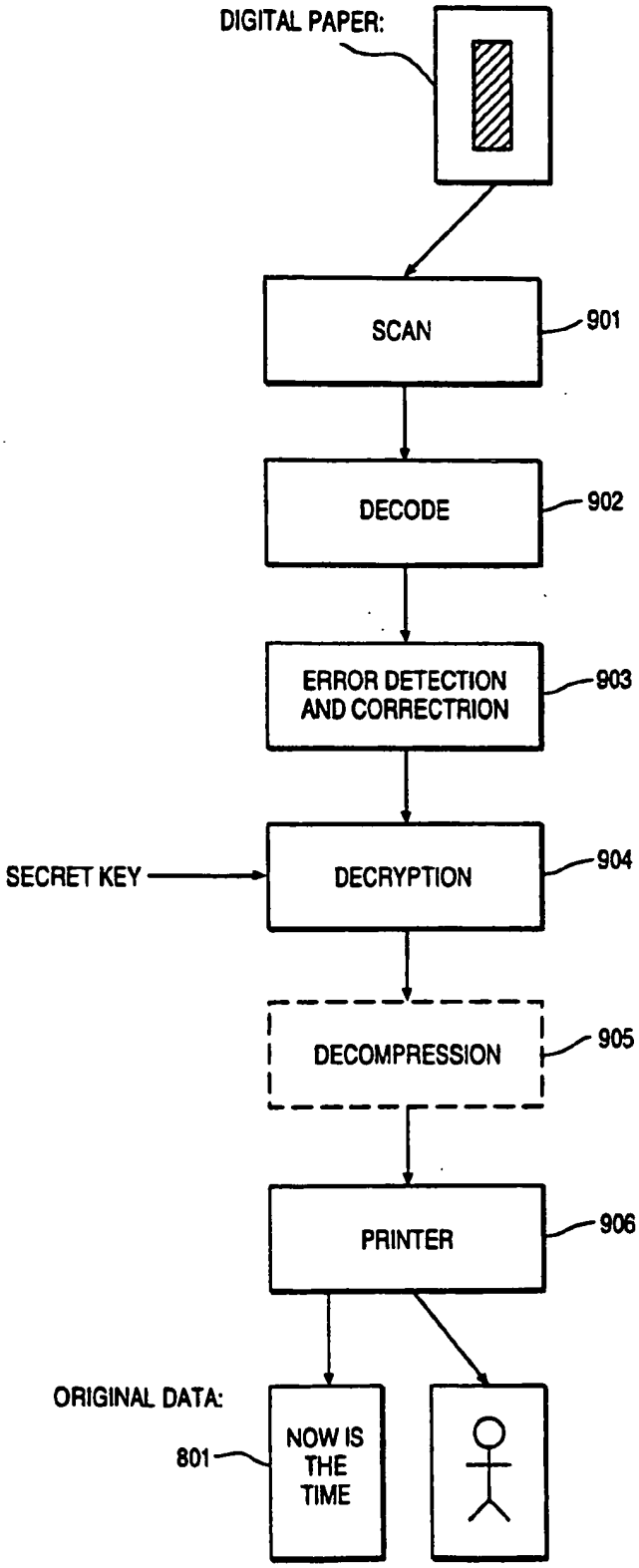


FIG_9 DECODING PROCESS



METHOD AND APPARATUS FOR PLACING DATA ONTO PLAIN PAPER

FIELD OF THE INVENTION

This invention relates to the field of digital communications; particularly, the present invention relates to digital communications on plain paper.

BACKGROUND OF THE INVENTION

Today, the most common practice of recording and transferring information is through the use of plain paper. One of the most widely recognized methods of recording information on plain paper is through the use of a photocopier. A photocopier scans an image, possibly performs some processing on the image and then reproduces the scanned image onto a piece of plain paper.

Until recently, prior art copiers mainly operated using analog signals. In other words, the operations of scanning the image, processing the image and then reproducing the image were all performed in analog. More recently, many functions of the copier have been performed in the digital arena. However, in either the analog or digital case, the function of the photocopier remains the reproduction of an original image onto a piece of plain paper.

The change from analog-to-digital signals in photocopiers allows several new operations to be performed. One new operation that may be performed is the encryption of information. Encryption is a technique whereby data may be stored in a format such that the data cannot be deciphered easily. In other words, encryption provides a mechanism for encoding data. There are numerous techniques in the prior art for encryption. However, all of these prior art techniques require that the data be digital. In the prior art, some copiers have been able to rearrange the information on the page in an attempt to provide encryption. One problem with providing encryption through simply rearranging the data on the page is that it isn't very secure. It would be advantageous to use a photocopier to perform encryption of digital data into plain paper such that it is secure.

In the prior art, some limited methods for storing digital information on plain paper have been set forth. One example of digital information being stored on paper is the use of bar codes. Bar codes are single dimension digital data in which typically ten digits of information are placed on a paper horizontally to store information. By scanning the bar codes in a relatively horizontal direction, the information encoded into the bar codes may be read. These scans usually occur very quickly. Note that the greater the height of the bar code (in the vertical direction), the less horizontal the scan is required to be to obtain the information. Also note that because bar codes are horizontal, there is an inherent limit to the number of bar codes which may be used to store digital information on a piece of paper. For more information on bar codes, see "Information Encoding with Two-Dimensional Bar Codes," authored by T. Pavlidis, J. Swartz and Y. Wang, COMPUTER, June 1992. See also U.S. Pat. No. 5,113,445 disclosing the placement of bar codes onto paper. It is desirable to store more digital information on a piece of plain paper than by using bar codes.

The present invention provides a method and apparatus for placing data on plain paper. In one embodiment, digital data is placed onto plain paper with a digital photocopier. The present invention also provides a method and apparatus for encrypting digital data such that the data remains secure upon reading the plain paper.

SUMMARY OF THE INVENTION

A method and apparatus for transferring digital information to and from plain paper is described. The present invention involves storing data in at least one box on plain paper. Each box includes a frame or border. Binary data is formatted in rows within the box, wherein a bit of digital data is depicted by the presence or absence of an ink dot. The inside edges of the left and right sides of the box have markers for identifying the vertical separations (i.e., the rows) between the digital data.

In one embodiment, the method and apparatus comprises a reproduction system having an encryption device for performing encryption on a data source to produce encrypted data. The present invention also includes an encoder for encoding the encrypted data into a series of pixel values. These pixel values are then output by a printing device, such that the data is placed in multiple boxes having a frame and a plurality of rows and columns on a piece of plain paper, such that the data source is represented in encrypted digitized form on the plain paper.

In the present invention, the reproduction system also includes a method and apparatus for recognizing the data on the plain paper. The method and apparatus for recognizing the data includes a scanning device for scanning the boxes on the plain paper and for converting the pixels into electrical signals representative of the characters. The present invention also includes a decoder for decoding the electrical signals into output signals representative of the data. A decryption device then performs decryption on the decoded data. An output device then transfers the decrypted data onto a piece of plain paper.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of the preferred embodiments of the invention which, however, should not be taken to limit the invention to the specific embodiments, but are for explanation and understanding only.

FIG. 1 is a block diagram of one embodiment of the encoding process of the present invention.

FIG. 2 is one embodiment of a block of encrypted digital data produced according to the present invention.

FIG. 3 illustrates a single sheet of digital paper of the present invention.

FIG. 4 is a block diagram of one embodiment of the decoding process of the present invention.

FIG. 5 is a block diagram of the digital processing system of the present invention.

FIG. 6 is a flowchart of a decoding routine of the decoding process of the present invention.

FIG. 7 illustrates one embodiment of the data box of the present invention being subjected to the decoding routine of the present invention.

FIG. 8 is a block diagram of one embodiment of the encoding process of the present invention that includes encryption.

FIG. 9 is a block diagram of one embodiment of the decoding process of the present invention that includes decryption.

DETAILED DESCRIPTION OF THE INVENTION

A method and apparatus for placing digital data onto plain paper is described. In the following description, numerous specific details are set forth such as processing steps, numbers of boxes, compression and encryption techniques, etc., in order to provide a thorough understanding of the present invention. It will be obvious, however, to one skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known operations have not been described in detail to avoid unnecessarily obscuring the present invention.

Overview of the Encoding Process of the Present Invention

FIG. 1 illustrates the encoding process of the present invention. The encoding process produces a sheet of paper which stores digital information. Thus, the result of the encoding process of the present invention is "digital" paper.

Referring to FIG. 1, in the present invention, data source **101** comprises digital data. In one embodiment, data source **101** is derived from a source of data that is scanned by a scanning device. In one embodiment, the scanning device is the scanner portion of a photocopier. Thus, in this embodiment, the data source **101** is a scan of paper on the photocopier glass. Note that in the currently preferred embodiment, the scan of the paper comprises digital data. Note also that the compression performed on the data has been signified in FIG. 1 by compression processing **102**. In other embodiments, data source **101** could come from an ASCII data file, a JPEG compressed color image or even a binary executable file. In other words, the data source **101** of the present invention can come from any source, including text, fax data, gray scale or even color image data. It should be noted that in the present invention, compression may or may not be performed. If compression is performed, various forms may be utilized, including both lossy and lossless compression techniques depending on whether the compressed data must be recreated in the exact form as the original.

Once the data source **101** is in digital form and any desired compression has taken place (processing block **102**), error correction encoding may be performed (processing block **103**). Since the plain paper is used as a digital channel, the error correction of the present invention can be viewed as just another box with digital input and digital output. Error correction encoding (processing block **103**) comprises adding correction or "parity" bits to the digital information of the encrypted data source **101**. In one embodiment, the error correction is performed using software. The software used to perform error correction is a shortened-interleaved Reed-Solomon code over GF (256) to provide the forward error correction. In the Reed-Solomon code over GF (256), one byte of data is taken from each box and combined using additions and multiplications in a finite field to produce two checksum bytes. In one embodiment, these bytes are encoded in the final two boxes on the page. For more information on Reed-Solomon Coding including the terms "shortened" and "interleaved", see Richard E. Blahut, *Theory and Practice of Error Control Codes*, Addison-Wesley, 1983.

As will be described below, in one embodiment, the data is stored on the plain paper in several boxes spread out on the page. The error correction code is interleaved by an amount equal to the number of bytes stored in one box on the page. It is therefore possible to lose an entire box without any uncorrectable errors. In this embodiment, one byte from each of the first ten boxes is used to compute the parity bits which are stored in the final two boxes. Note that in the present invention the boxes with the parity bits are no different than the other boxes of digital data on the plain paper. One parity box can be lost and the entire message can still be constructed just as for the data boxes. In one embodiment, twelve boxes of data are used for data storage on the paper.

Note that there is an advantage to using information about the formatter to select the parameters the error correction. That is, since error correction can group any set of bytes together, if the location of the bytes on the page is known, then bytes used in the error corrections may be those which are separated on the page. In this manner, error correction results which can better handle situations where a large portion of the plain paper has sustained damage (for instance, such as the damage due to a coffee spill). In addition, parameters can be chosen based on the decoder. Thus, if a very high resolution scanner is used in decoding, then the coder could attempt to transmit more data. If the scanner has a large platen, then more information can be decoded (all other things being equal). If it is known that the particular receiver is highly noisy, then a more robust error correction method could be used.

Once the data source **101** is in digital form, any desired compression has occurred (processing block **102**), and any desired error correction encoding has been performed (processing block **103**), the digital data is formatted into distinct colored pixels (processing block **104**) which can be printed. In one embodiment, the digital data is formatted into black and white pixels. In one embodiment, the pixels are printed by the printer portion of the photocopier.

The formatting step of the present invention (processing block **104**) is responsible for changing a digital bit sequence into a form which can be accurately scanned and a format which is readable by the printer. In one embodiment, the formatter produces a Postscript TM file which can be printed on a laser printer.

After formatting (processing block **104**), the data is printed (processing block **105**). In one embodiment, where the formatting process produces a Postscript file describing the page of the formatted data, the Postscript file is then printed on a high quality Postscript printer, such as the LaserWriter IINT brand printer of Apple Corporation of Cupertino, Calif.

In order to write the bits on the paper, initially a stream of data to be stored on the piece of paper is stored as a block of black and white dots, or data pixels, on the paper. In one embodiment, a "1" bit can be stored as a white dot and "0" bit as a black dot. The data pixels utilized must be larger than the scanning resolution so that the scanning sample will lie within the printed pixel. In addition, a frame (or border)is placed around the block of data so that the position to read each pixel from can be accurately determined. An example of one such digital data box is shown in FIG. 2. If the data block is too large, problems with paper

stretch and scanner misalignment may make reconstruction of the original data difficult.

Referring to FIG. 2, in the currently preferred embodiment, to ensure the data spacing and alignment could be determined, the black and white bits are placed in several boxes on the page. In one embodiment, twelve boxes of data are used for each $8\frac{1}{2} \times 11$ inch sheet of paper. In one embodiment, each box has a border that is three data pixels thick and mostly black. By having a border around each box, it is easier to locate the edge of the box (when decoding). Note also that in each corner of each box, there is a large white pixel. An examination of the box also reveals alternating pixels along both the left and right edges. These are used to accurately determine the current location of a horizontal line of pixels when reading the data. Note that the alternating pattern of black and white pixels is added to the left and right edges of each box to compensate for the possible variance in the vertical spacing of the pixels. These alternating pixels are not required where the vertical spacing between the pixel rows does not vary by an amount which would cause the incorrect reading of the data (upon decoding).

Note that although twelve small boxes each storing several bytes of data is used in one embodiment, any number of boxes could be used, including only a single box. This is a design choice. A small box uses more overhead for the frame, and the space between the boxes is unused for data, but requires less correction data to be added to the encoded data (i.e., a lower error rate).

Note also that the present invention may use gray-scale inks and store additional information in the gray level. In this manner, for example, eight shades of gray could be used to obtain three bits of data from a gray-scale printer and scanner.

Overview of the Decoding Process of the Present Invention

FIG. 4 illustrates the decoding process of the present invention. The decoding process of the present invention inverts the operation of the encoding process one step at a time. Initially, the sheet of digital paper is scanned (processing block **401**). The scanned data is then sampled to produce a binary sequence representing the one received by the encoding process (processing block **402**). Note that the result of the scan is a binary sequence which is close to the one given to the encoder. The result of the scanning may not produce an exact duplicate of the data sent to the encoder. This may be due to errors in the scanning of the data. For instance, an exact duplicate may not be sent to the decoder due to errors in the registration, flatness, skew, or scanner defects during the scanning process.

After the "digital" paper has been decoded (processing block **402**), error detection and correction is performed (processing block **403**). The error correction compensates for the loss of data due to damage to the paper or due to the failure to predict the pixel location correctly. The error detection and correction processing (processing block **403**) is the inverse of the error correction encoding which occurred during the encoding process (FIG. **1**). The error detection and correction occurs by using the parity bits added during the encoding process of FIG. 1 to correct errors which may have occurred (processing block **403**). The errors may be attributed to staples, hole punches, paper discolorations, technical problems or other paper imperfections and damages.

Once any required error correction is performed, the data may undergo decompression (processing block **404**). The decompression performed depends on whether compression was performed during the encoding process and is performed to restore the data, as near as possible, to its original form.

After decoding (processing block **402**), error detection and correction (processing block **403**) and any desired decompression (processing block **404**), the reconstructed data is available and can be printed, displayed or stored on a disk. In one embodiment, the data is printed (processing block **405**). In one embodiment, if the original data was a simple scan, a copy of the original can be printed. In another embodiment, if the original data was several pages of ASCII or Postscript data, the data can be printed or stored in a file. In another embodiment, if the original was a JPEG compressed color image, the image may be displayed.

One Embodiment of the Encoding/Decoding System of the Present Invention

FIG. 5 illustrates an overview of one embodiment of the encoding and decoding system of the present invention shown in block diagram form. The system of the present invention is a digital processing system. In the currently preferred embodiment, the digital processing system comprises a digital photocopier. In one embodiment, the operation of a digital photocopier is simulated using a scanner, printer and computer. It will be understood that while FIG. 5 is useful for providing an overall description of the processing system of the present invention, a number of details of the system are not shown. As necessary for disclosure of the present invention, further detail is set forth with other figures provided with this specification. Further, the present invention is described with reference to its preferred embodiment; alternative embodiments which may be conceived by one of ordinary skill in the art are considered within the scope of the claims set forth below.

Referring to FIG. 5, the system of the present invention includes a bus or other communication means **501** for communicating information. Processor **502** is coupled with bus **501** for processing information. A random access memory (RAM) or other dynamic storage device **503** (commonly referred to as a main memory) for storing information and instructions for processor **502** is also coupled to bus **501**. Also coupled to bus **502** is a read only memory (ROM) or other static storage device **504** for storing static information and instructions for processor **102** and a data storage device **505**, such as a magnetic disk and disk drive for storing information and instructions.

The processing system also includes a scanner **506** coupled to bus **501** for scanning selected hard copy documents into the processor. Scanner **506** is capable of reading digital representations of images (i.e., digital paper), as well as regular images. In order to differentiate between whether scanner **506** is reading a regular image or a piece of digital paper, some form of automation could be employed to search a portion of the document to determine that the hard copy document being scanned is a piece of digital paper. In another embodiment, a key could be entered which indicates to scanner **506** that the hard copy being scanned is a piece of digital paper.

In one embodiment, scanner **506** comprises a gray scale scanner. In the currently preferred embodiment, this resolution is 200 DPI. Scanner **506** converts the individual picture elements, referred to as pixels, of the scanned image into digital values. In one embodiment, scanner **506** comprises an Image Scanner ICS-400 brand scanner of Ricoh Corporation of West Caldwell, N.J. In other embodiments, scanner **506** is a bitmap scanner which scans the image of each hard copy input document in a predetermined spatial resolution to produce digital values. These digital values collectively produce a data structure known as a bitmap image, which is well-known to those in the art.

Note that processor **502** performs the error correction encoding, encoding, decoding, and error detection and correction processes of FIGS. **1** and **4** in conjunction with the input from scanner **506**. Note also that processor **502** may also perform encryption and decryption if employed in the digital paper process. Thus, in one embodiment, processor **502** operates as the encoder and the decoder of the present invention. In one embodiment, processor **502** comprises a SPARC station **2** brand processor of Sun Microsystems, Inc. of Mountain View, Calif.

A hard copy device **507** is also coupled to bus **501** for printing hard copies for providing visual representations of the scanned originals. In one embodiment, hard copy device **507** comprises the printer portion of a photocopier. In other embodiments, hard copy device **507** could comprise a plotter or printer, such as a bitmap printer that maps the digital values of a bitmap image into pixels which are printed on the plain paper.

Moreover, a human or user interface **508** is included for enabling a user to interact with processor **502**, scanner **506** and hard copy device **507**. User interface **508** represents the input and output devices through which the user enters control instructions to and receives feedback from the photocopier (i.e., processor **502**, scanner **506** and hard copy device **507**). The feedback indicates the actions that are taken by the photocopier in response to instructions that are entered by the user. User interface **508** may be used to enter a key for encryption and decryption processes, as will be described in conjunction with FIGS. **8** and **9** below.

User interface **508** could include an alphanumeric input device including alphanumeric and other keys for communicating information and command selections, a cursor control device for controlling cursor movement and/or a display device, such as a cathode ray tube, liquid crystal display, etc. for displaying information to the photocopier user. Note that these components are well-known in the art and have been omitted to avoid unnecessarily obscuring the present invention. Note that in one embodiment, user interface **508** may provide inputs directly to scanner **506** and hard copy device **507** which can be distinct from those inputs to processor **502**.

Note that, although in the currently preferred embodiment, all of these components are integrated into a single photocopier system, each of the components, such as scanner **506** and hard copy device **507**, may be separate components. In one embodiment, scanner **506** and/or hard copy device **507** could be coupled to bus **502** using dedicated communication links or switchable communication networks.

Scanning Digital Paper

In order to decode the digital paper, the entire sheet is scanned with a scanner at high resolution. In one embodiment, the entire sheet is scanned with a gray scale scanner. A software routine is used to generate a binary sequence from the scanned in data. The result produced by the software routine is a binary sequence. A flow chart of the decoding routine is shown in FIG. **6**. Note that FIG. **7** illustrates a portion of the decoding process.

Referring to FIG. **6**, once the digital paper is scanned in an over sampled manner, the decoding routine begins by locating the boxes on the page by searching for the black border (processing block **601**). In one embodiment, the search for the black border is accomplished by scanning a row until a predetermined number of black pixels are found in a row. In the currently preferred embodiment, the borders are 3 pixels thick. The number and approximate position of the boxes on the page is assumed to be known by the decoding process (since how the digital paper is produced is known by the decoding process). However, this is not a requirement since the frame should be easy to identify even on a page full of text. Note also that in one embodiment, the size of the boxes is known to the decoding process. Therefore, once a border has been located, the other borders are located by searching the vertical and horizontal directions until the other borders of the box are located (processing block **602**).

After the borders of the data blocks have been located, the four white dots are located at each corner of each box (processing block **603**). In one embodiment, these white dots are located by searching the corners for the brightest pixel. Since the scanning resolution used in one embodiment is at least twice the printing resolution, each corner should have several white scanned pixels.

Because an encrypted page may be copied several times, the spacing between information pixels may change and the rows of pixels will not be perfectly aligned with the scanning row. The brightest pixel is used as the correct position to find the remainder of the data. In other words, the four white corner dots are used by the present invention to determine the horizontal spacing between the information pixels within the boxes. Once the width and height of the box (in number of data pixels) are known and the corners have been located, the horizontal and vertical spacing is determined (processing block **604**).

To find the horizontal and vertical spacing when the corners have been found, the alternating pattern of black and white pixels added to the right and left edges of the box during the encoding process to compensate for the variance of the vertical spacing are examined to determine the vertical spacing between the information pixels. First, the top two corners are joined by an imaginary line (processing block **604A**). The joining of the two top corners is explicitly shown in FIG. **7**. A synchronization column is then approximated as a fraction of a distance between the corners (processing block **604B**). The synchronization column has been shown as an imaginary line in FIG. **7**. Next, the decoder finds all of the transitions in the synchronization column (processing block **604C**). In one embodiment, the decoder records the location of every black-to-white or white-to-black transition. Then, starting from the point between two transitions on the left, an imaginary line can

be drawn to the point between two transitions on the right (processing block 604D). Note that these lines are drawn in FIG. 7. Even though one of the rows of data is much shorter than the others and one is much taller, the correct place to read the rows can still be determined. Finally, since the number of data pixels is known, this horizontal line can be divided into the correct number of pieces and sampled at the correct location (processing block 604E).

With the vertical and horizontal spacing determined, it is possible to compute the two dimensional grids from which the data pixels can be sampled. Finally, the data box is examined horizontally and if the sample at the predicted location is closer to black than white, a 0-bit is emitted; otherwise, a 1-bit is emitted (processing block 605). In one embodiment, a bi-linear interpolation of the four pixels closest to the desired location is used to determine the value of the data. Given an interpolation of the value of the pixel at the real location, the average is compared with expected values for white or black pixels. Then if the sampled pixel is closer to white, the pixel is assumed to be a "1" data bit; otherwise, the data is assumed to be a "0" data bit. Note that other methods of interpolating and sampling may be used and are known to those skilled in the art. For example, a faster method would use the pixel closest to the desired sample point without any interpolation.

Note that encoding digital paper is similar to disk storage because binary data is mapped into space. Therefore, well-known disk storage systems, such as RLL, MFM, M^2FM and others, can be used to improve encoding. Most of these methods utilize data transitions rather than data states. Data would then be encoded by the location of a change from one color to another, rather than the color itself. Another possible improvement is using the ink transitions to accurately determine the pixel spacing and data locations. The transitions thus provide not only data, but also the pixel spacing. Decoding is done with a phase lock loop algorithm.

Pixel Size and Paper Curl

In the present invention, the size of the pixel is influenced by two considerations. First, the size of the pixel is influenced by the amount of information that is to be placed onto the paper. The smaller the size pixel, the greater the amount of information which may be placed onto the paper. Second, the size of the pixel is influenced by the capability of the scanner used to scan the data from the paper when decoding. The smaller the pixel, the higher the scan rate that is required.

In one embodiment, the data pixels are printed at 50 DPI, while the scanning is performed at 200 DPI. This allows fast decoding and error correction which results in zero errors being observed in the digital data. Printing of data pixels can occur at 100 DPI and scanning at 400 DPI to increase the storage on a single sheet of paper from 15 kilobytes to 60 kilobytes.

Note that the current software uses a parameter file to set the printing and scanning resolutions and box sizes. In this manner, changes to each of these parameters may be made easily. It is possible to print at 50 DPI and scan at 200 DPI without errors. It is also possible to print at 100 DPI and scan it, 400 DPI without errors. This would allow more than 60 kilobytes of data per page to be stored. If absolutely error free operation is not required, a wide variety trade-offs between printing size, scanning size, amount of error correction, and error percentage is possible. In general, if errors are allowed, the scanning resolution can be much closer to the printing resolution and much more data can be stored on a page.

Normal scanner/photocopier covers keep paper from shifting away from the scanner lens and prevents bad lighting from occurring during the scanning process. Note that during the scanning process, scanning errors can be reduced by ensuring that the plain paper is kept flat enough to maintain accurate pixel spacing. There are numerous methods and devices which may be used to keep the paper flat and in place during scanning.

Thus, the present invention provides a method and apparatus for storing and transferring digital data on an ordinary sheet of paper, as well as retrieving the data through the use of a digital scanner. Neither the encoding nor the decoding of the data require very much processing power. The present invention allows the storage of more than 15 kilobytes of information virtually error free. The present invention also provides a method of digital transmission which allows more data to be stored on a page than text currently does. A typical page or text may have forty lines of eighty characters or 3,200 total characters. If each character truly represented eight bits of information, then one page would hold three kilobytes. Use of the present invention results in the storage of more than 15 kilobytes of data on a page. If the original text is English, it can be compressed by a factor of three and then one sheet of digital paper could be used to store the equivalent of 15 pages of text. Also, it is possible to store a normal color image on a black and white copy by storing the information digitally. Even more data could be stored with the previously mentioned disk storage codes.

Furthermore, the present invention allows for digital data to be printed and read by off-the-shelf photocopier products. In other words, the present invention does not require special laser scanners or hand-held wands to transfer digital data to and from a piece of plain paper.

Encryption

In one embodiment, encryption and decryption may be integrated into the encoding and decoding processes of the present invention. Note that in one embodiment of the present invention, this would occur after the compression of the digital data (processing block 102 of FIG. 1). One embodiment of the encoding process that includes encryption is shown in FIG. 8. Note that each of the processing blocks in FIG. 8 are performed as their similarly named counterparts in FIG. 1, with the exception of encryption processing block 803.

In one embodiment, the encryption is accomplished using software. In one embodiment, the software encryption is provided by exclusive-ORing the initial data with a pseudo-random sequence generated using a secret key as a seed. Note that in this case, the decryption process (as described below later)is simply exclusive-ORing the data with the same pseudo-random sequence. Note that this type of encryption process is not perfectly secure method of data encryption because of the short key length used to seed pseudo-random sequences and the fact that the same key is likely to be used on more than one message. There are well-known methods to attack this type of encryption.

In order to obtain secure data through encryption, a one time pad can be employed in place of the pseudo-random sequence. A one time pad comprises a series of randomly generated bits that are known to both parties (i.e., the "encrypting" party and the "decrypting"

party) and is only used once in the encryption process. In a one time pad, the key length is equal to the message length and may be more than 100,000 bits. Since the one time pad is used once, it is not susceptible to the same attacks which are used on pseudo-random sequences.

In another embodiment, encryption processing could be performed using the digital encryption standard (DES) or the RSA algorithm. In the RSA algorithm, digital data is encrypted using two prime numbers which are multiplied together, as is well-known in the art. In the case where each of the two prime numbers has approximately 200 digits, the RSA algorithm offers a very secure encryption method. Note that to use DES in one of the feedback modes or to use the RSA algorithm, it is necessary that errors do not occur in the encoding and decoding of the digital data stream. If an error occurs, any information beyond the location of the error would be lost. Alternatively, the encryption process can periodically be restarted so that only a portion of the message is lost if too many errors occur. It should be noted that any digital encryption method may be employed with the present invention.

Similarly, if encryption is integrated into the encoding process, then decryption must be included in the decoding process. One embodiment of the decoding process of the present invention that includes decryption is shown in FIG. 9. Note that each of the processing blocks in FIG. 9 are performed in the same manner as their similarly named counterparts in FIG. 4, with the exception of decryption processing block 904.

After error detection and correction (processing block 903), the data undergoes decryption (processing block 904). The decryption processing (processing block 904)is the inverse of the encryption performed during the encoding process (FIG. 8). In one embodiment, the decryption process requires the use of the same key used during encryption. By using the same key, the original data is recreated.

When the methods of encryption (and decryption) of the present invention are being integrated into the encoding and decoding processes, the present invention also provides data information to be transferred using plain paper in a manner which preserves its privacy and/or authentication. In the present invention, this privacy can easily be obtained through the use of a key or code known to the user(s) when encryption and decryption of a document occurs. Note that the present invention provides for encryption of information using the photocopier. Using the present invention, a user could place a confidential document on a copier, press an encrypt button and enter a secret key. The copier would then scan the sheet and produce an encrypted copy. This encrypted copy could be mailed or otherwise transported to the intended receiver, who places it on a copier, presses a decryption button and enters the same secret key. The copier then produces a readable copy of the original document. The present invention also allows the encrypted copy to be treated as a normal document and be copied or stapled without greatly affecting the final decrypted document. The present invention also provides for secure encryption of a document because any person who does not know the secret key will be unable to determine the contents of the encrypted document.

Another useful application of the present invention is the authentication of a document. Specifically, the present invention could be used to authenticate signed facsimile transmissions. Through the use of digital signatures, authentication could be simple. If a photocopier allowed digital signing of a document, then digital signatures could be used for all business transactions. The authentication might consist of a small block of digital, encrypted data on an otherwise human readable document. Using the secret key, a person would be able to obtain the digital data in the small block, therein being able to verify its source. In addition, the codes could prevent forgery or provide secret information since only those using the correct key or code would be able produce the digital paper.

Whereas many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after having read the foregoing description, it is to be understood that the particular embodiment shown and described by way of illustration are in no way intended to be considered limiting. Therefore, reference to the details of the preferred embodiments are not intended to limit the scope of the claims which themselves recite only those features regarded as essential to the invention.

Thus, a method and apparatus for placing digital information on plain paper has been described.

I claim:

1. A system for placing a data source on plain paper as digital data comprising:

formatting means for formatting the data source into a series of digital data values, wherein the series of digital data values are formatted into a plurality of contiguous rows and a plurality of contiguous columns of pixels in at least one data box, wherein each pixel represents at least one bit in the series of digital data values, and further wherein the data box includes a frame bordering the series of digital data values having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box; and

printing means coupled to the formatting means for printing said at least one box onto at least one piece of plain paper having the plurality of rows and the plurality of columns, such that the data source is represented in digitized form on the plain paper.

2. The system as defined in claim 1 wherein the formatting means formats the data into a series of binary values.

3. The system as defined in claim 1 wherein a plurality of boxes having data values are printed on the plain paper.

4. The system as defined in claim 1 wherein a bit of data is printed on the plain paper, such that the presence of a dot on the plain paper indicates that the bit is in a first logical state and the absence of a dot on the plain paper indicates that the bit is in a second logical state.

5. The system as defined in claim 1 wherein a bit of data is printed on the plain paper, such that the change between the presence or absence of a dot on the plain paper indicates whether a bit is in a first logical state or is in a second logical state.

6. The system as defined in claim 1 wherein each pixel value printed represents three bits of data.

7. The system as defined in claim 1 further comprising compression means for compressing the data source before the formatting means formats the data.

8. A system for placing a data source on plain paper as digital data comprising:

formatting means for formatting the data source into a series of binary values, wherein the series of binary values are formatted into a plurality of contiguous rows and a plurality of contiguous columns of pixels in a plurality of data boxes, wherein each pixel represents at least one bit in the series of binary values, and further wherein each of the plurality of data boxes includes a frame bordering the series of binary values having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels; and

printing means coupled to the formatting means for printing said plurality of data boxes onto at least one piece of plain paper, wherein the presence of a dot on the plain paper indicates that the bit is in a first logical state and the absence of a dot on the plain paper indicates that the bit is in a second logical state, wherein the plurality of contiguous rows and a plurality of contiguous columns of pixels are printed in the plurality of rows and the plurality of columns in the plurality of data boxes, such that the data source is represented in digitized form on the plain paper.

9. The system as defined in claim 8 further comprising compression means for compressing the data source before the formatting means formats the data.

10. The system as defined in claim 8 further comprising means for error correction.

11. The system as defined in claim 8 wherein each pixel value printed represents a plurality bits of data.

12. A system for decoding plain paper having at least one data box for storing binary data values in the form of pixels, wherein said at least one data box includes a plurality of borders, said system comprising:

first locating means for locating means for locating said at least one data box on the plain paper, wherein said first locating means includes means for scanning rows of the plain paper until a predetermined number of contiguous pixels of a first color are located in a row, such that said first locating means locate one of the plurality of borders, said first locating means further comprising means for locating the other of the plurality of borders;

second locating means for locating pixels of a second color in corners of the plurality of borders of said at least one data box;

third locating means for locating the plurality of rows of pixels in each said at least one data box on the plain paper;

decoding means coupled to the first locating means, the second locating means and the third locating means for decoding each pixel in the plurality of rows into a binary data value, wherein the presence of a dot on the plain paper indicates that the bit is in a first logical state and the absence of a dot on the plain paper indicates that the bit is in a second logical state, and further wherein the decoding means locates pixels in said at least one data box based on horizontal spacing of pixels determined by the second locating means and vertical spacing of pixels determined by the third locating means, such that the plain paper is decoded.

13. The system as defined in claim 12 wherein said at least one data box includes at least two opposite edges having markers for identifying the vertical separations between the rows of data in said each of the plurality of data boxes, such that said third locating means locates each of the plurality of rows of binary data using said at least two opposite edges.

14. A method for placing digital data on plain paper comprising the steps of:

formatting the data source into a series of binary values, wherein the series of binary values are formatted into a plurality of contiguous rows and a plurality of contiguous columns of pixels in a plurality of data boxes, wherein each pixel represents at least one bit in the series of binary values, and further wherein each of the plurality of data boxes includes a frame bordering the series of binary values having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of pixels and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box; and

printing the plurality of data boxes onto at least one piece of plain paper, wherein each of the plurality of binary values is either printed or not printed according to its logic state, such that the presence of a dot on the plain paper indicates that the bit is in a first logical state and the absence of a dot on the plain paper indicates that the bit is in a second logical state, wherein the binary values are printed in the plurality of rows and the plurality of columns in the plurality of boxes, such that the data source is represented in digitized form on the plain paper.

15. A method for decoding plain paper having at least one data box for storing binary data values in the form of pixels, said method comprising the steps of:

first locating said at least one data box on the plain paper, including the step of scanning rows of the plain paper until a predetermined number of contiguous pixels of a first color are located in a row, such that one of a plurality of borders of said at least one data box is located, and including the step of locating the other of the plurality of borders;

second location pixels of a second color in corners of the plurality of borders of said at least one data box, wherein horizontal spacing of pixels in said at least one data box is determined:

third Locating the plurality of rows of pixels in each of said at least one data box on the plain paper, wherein vertical spacing of pixels in said at least one data box is determined;

decoding each pixel in the plurality of rows into a binary data value, wherein the presence of a dot on the plain paper indicates that the bit is in a first logical state and the absence of a dot on the plain paper indicates that the bit is in a second logical state, and further wherein the pixels in said at least one data box are located based on horizontal spacing of pixels determined by the second locating step and vertical spacing of pixels determined by the third locating step, such that the plain paper is decoded.

16. A reproduction system for reproducing a data source comprising:

encryption means for performing encryption on the data source to produce encrypted data representing the data source, wherein the data source is digital and the encrypted data produced is digital;

encoding means coupled to the encryption means for encoding the encrypted data into a series of pixel values; and

printing means coupled to the encoding means for printing the series of pixel values onto at least one piece of plain paper, wherein the pixel values are printed in at least one box in a plurality of contiguous rows and a plurality of contiguous columns of pixels, wherein each pixel represents at least one bit in the series of digital data values, and wherein said at least one box includes a frame bordering the series of pixels having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box, such that the data source represented in encrypted digitized form on the plain paper.

17. The reproduction system as defined in claim 16 further comprising scanning means for scanning the data source into digitized signals.

18. The reproduction system as defined in claim 16 further comprising means for performing error correction encoding on the encrypted data.

19. The reproduction system as defined in claim 16 wherein the encryption means uses a digital key.

20. The reproduction system as defined in claim 19 wherein the key seeds a pseudo-random sequence.

21. The reproduction system as defined in claim 19 wherein the key comprises a one-time pad.

22. The reproduction system as defined in claim 16 wherein the encryption means includes means for encrypting the data according to the digital encryption standard (DES).

23. The reproduction system as defined in claim 16 wherein the encryption means includes means for encrypting the data according to the RSA algorithm.

24. A system for representing and recognizing data on plain paper in the form of a plurality of data boxes comprising:

encoding means for encoding the data, wherein the encoding means includes:

encryption means for performing encryption on the data source to produce encrypted data representing the data source, wherein the data source is digital and the encrypted data produced is digital;

encoding means coupled to the encryption means for encoding the encrypted data into a series of pixel values; and

first output means coupled to the encoding means for outputting the series of pixel values onto at least one piece of plain paper, wherein the series of pixel values are printed in at least one box in a plurality of contiguous rows and a plurality of contiguous columns and wherein said at least one box includes a frame bordering the series of pixels having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box, such that the data source represented in encrypted digitized form on the plain paper; and

recognition means for recognizing the data in the plurality of boxes, wherein the recognition means includes:

scanning means for scanning said at least one box on the plain paper and for converting the pixels into electrical signals representative of the characters, wherein the scanning means scans rows of the plain paper until a predetermined number of contiguous pixels of a first color are located in a row, such that one of a plurality of borders of said at least one data box is located, and further wherein said scanning means locates the other of the plurality of borders, said scanning means also including means for determining the horizontal spacing and the vertical spacing of pixels in said at least one data box;

decoding means coupled to the scanning means for decoding the electrical signals into output signals representative of the data;

decryption means coupled to the decoding means for decrypting the electrical signals; and

second output means coupled to the decrypting means for transferring the output signals onto a piece of plain paper.

25. The reproduction system as defined in claim 24 wherein the encryption means uses a digital key.

26. The reproduction system as defined in claim 25 wherein the key seeds a pseudo-random sequence.

27. The reproduction system as defined in claim 25 wherein the key comprises a one-time pad.

28. A process for reproducing a data source comprising the steps of:

performing encryption on the data source to produce encrypted data representing the data source, wherein the data source is digital and the encrypted data produced is digital;

encoding the encrypted data into a series of pixel values; and

outputting the series of pixel values onto at least one piece of plain paper, wherein the pixel values are printed in at least one box in a plurality of contiguous rows and a plurality of contiguous columns and wherein said at least one box includes a frame bordering the series of pixels having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box, such that the data source represented in encrypted digitized form on the plain paper.

29. A process for decoding the data in the plurality of boxes comprising the steps of:

scanning said at least one box on the plain paper and for converting the pixels into electrical signals representative of the characters, wherein each of the pixels; represents a binary value;

decoding the electrical signals into output signals representative of the data, wherein the step of decoding includes the steps of:

locating the corners of said at least one box using pixels in the corners;

determining the spacing of the pixels according to the location of the corners of said at least one box, wherein horizontal spacing of the pixels is determined using the pixels in the corners of said at least one box;

locating the rows of pixels in said at least one box using alternating pixels along two opposite sides of said at least one box, such that the vertical separation between the plurality of rows of data is identified; and

sampling the pixels according to the horizontal spacing and the vertical separation in order to produce decoded data;

decrypting the decoded data; and

transferring the output signals onto a piece of plain paper.

30. The process as defined in claim 29 wherein the step of decrypting the decoded data comprises the step of decrypting the decoded data according to a key.

31. The process as defined in claim 29 wherein the step of decrypting the decoded data comprises the step of decrypting the decoded data according to a key seeding a pseudo-random sequence.

32. The process as defined in claim 29 wherein the step of decrypting the decoded data comprises the step of decrypting the decoded data according to a key comprising a one-time pad.

33. A system for placing a data source on plain paper as digital data comprising:

a formatter, wherein the formatter formats the data source into a series of digital data values, wherein the series of digital data values are formatted into a plurality of contiguous rows and a plurality of contiguous columns of pixels in at least one data box, wherein each pixel represents at least one bit in the series of digital data values, and further wherein the data box includes a frame bordering the series of digital data values having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box; and

a printer coupled to the formatter, wherein the printer prints said at least one box onto at least one piece of plain paper having the plurality of rows and the plurality of columns, such that the data source is represented in digitized form on the plain paper.

34. The system defined in claim 33 further comprising a data compresser coupled to receive the data source, wherein the data compresser compresses the data source and supplies the data source in a compressed form to the formatter for formatting.

35. The system defined in claim 33 further comprising an error correction encoder coupled to receive the data source, wherein the error correction encoder performs error correction encoding on the data source and supplies the data source in an error correction encoded form to the formatter for formatting.

36. A system for placing a data source on plain paper as digital data comprising:

a data compresser coupled to receive the data source, wherein the data compresser compresses the data source to produce a compressed data source;

an error correction encoder coupled to receive the compressed data source, wherein the error correction encoder performs error correction encoding on the compressed data source to produce a compressed and error correction encoded data source;

a formatter coupled to receive the compressed and error correction encoded data source, wherein the formatter formats the compressed and error correction encoded data source into a series of digital data values, wherein the series of digital data values are formatted into a plurality of contiguous rows and a plurality of contiguous columns of pixels in at least one data box, wherein each pixel represents at least one bit in the series of digital data values, and further wherein the data box includes a frame bordering the series of digital data values having alternating pixels along at least two opposite sides for identifying the vertical separations between the plurality of rows of data and having pixels in corners in the frame to determine horizontal spacing between pixels within said at least one data box; and

a printer coupled to the formatter, wherein the printer prints said at least one box onto at least one piece of plain paper having the plurality of rows and the plurality of columns, such that the data source is represented in digitized form on the plain paper.

37. The system defined in claim 3 wherein at least one of the plurality of boxes contains error correction encoded data for the other of the plurality of boxes.

* * * * *

US005781914A

United States Patent [19]

Stork et al.

[11] Patent Number: 5,781,914

[45] Date of Patent: Jul. 14, 1998

[54] **CONVERTING DOCUMENTS, WITH LINKS TO OTHER ELECTRONIC INFORMATION, BETWEEN HARDCOPY AND ELECTRONIC FORMATS**

[75] Inventors: **David G. Stork**, Stanford; **K. Venkatesh Prasad**, Cupertino, both of Calif.

[73] Assignees: **Ricoh Company, Ltd.**, Tokyo, Japan; **Ricoh Corporation**, Menlo Park, Calif.

[21] Appl. No.: **497,985**

[22] Filed: **Jun. 30, 1995**

[51] **Int. Cl.**[6] **G06F 17/40**

[52] **U.S. Cl.** **707/506**; 707/501; 707/513

[58] **Field of Search** 395/761, 762, 395/774; 382/175; 358/403; 707/501, 505–508, 513–516, 524, 530

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,113,445	5/1992	Wang	380/51
5,317,646	5/1994	Sang, Jr. et al.	707/506
5,333,237	7/1994	Stefanopolis et al.	707/501
5,337,362	8/1994	Gormish et al.	380/54
5,367,621	11/1994	Cohen et al.	707/501
5,486,686	1/1996	Zdybel, Jr. et al.	307/541 X
5,557,722	9/1996	DeRose et al.	707/513

OTHER PUBLICATIONS

Myka et al., Using Electronic Facsimile of Documents for Automatic Reconstruction of Underlying Hypertext Structures. Document Analysis, pp. 528–532, Oct. 20, 1993.

Iizawa et al., Avital, A Private Teaching System by FAX Communication, Systems Integration 2nd Intl Conf., pp. 293–301, Jun. 15, 1992.

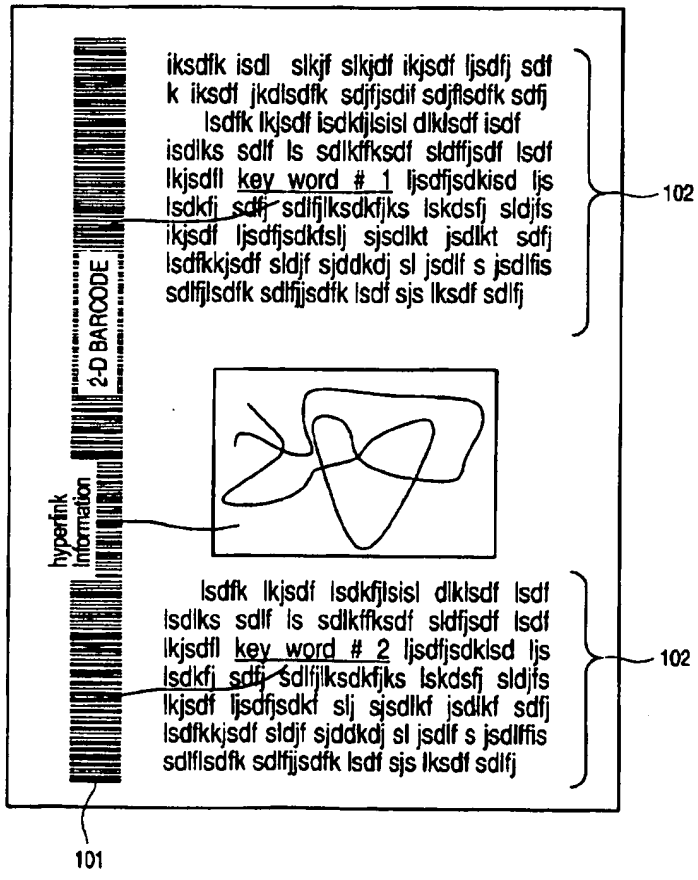
Primary Examiner—Anton Fetting

Attorney, Agent, or Firm—Blakely, Sokoloff, Taylor & Zafman

[57] **ABSTRACT**

A conversion method and apparatus that allows for converting a hardcopy document into a hyperdocument and vice versa. During hardcopy to hyperdocument conversion, hypertext information stored on the hardcopy document is used to set up links to other documents. During hyperdocument to hardcopy document conversion, hypertext link information is encoded and stored on the hardcopy document.

23 Claims, 10 Drawing Sheets



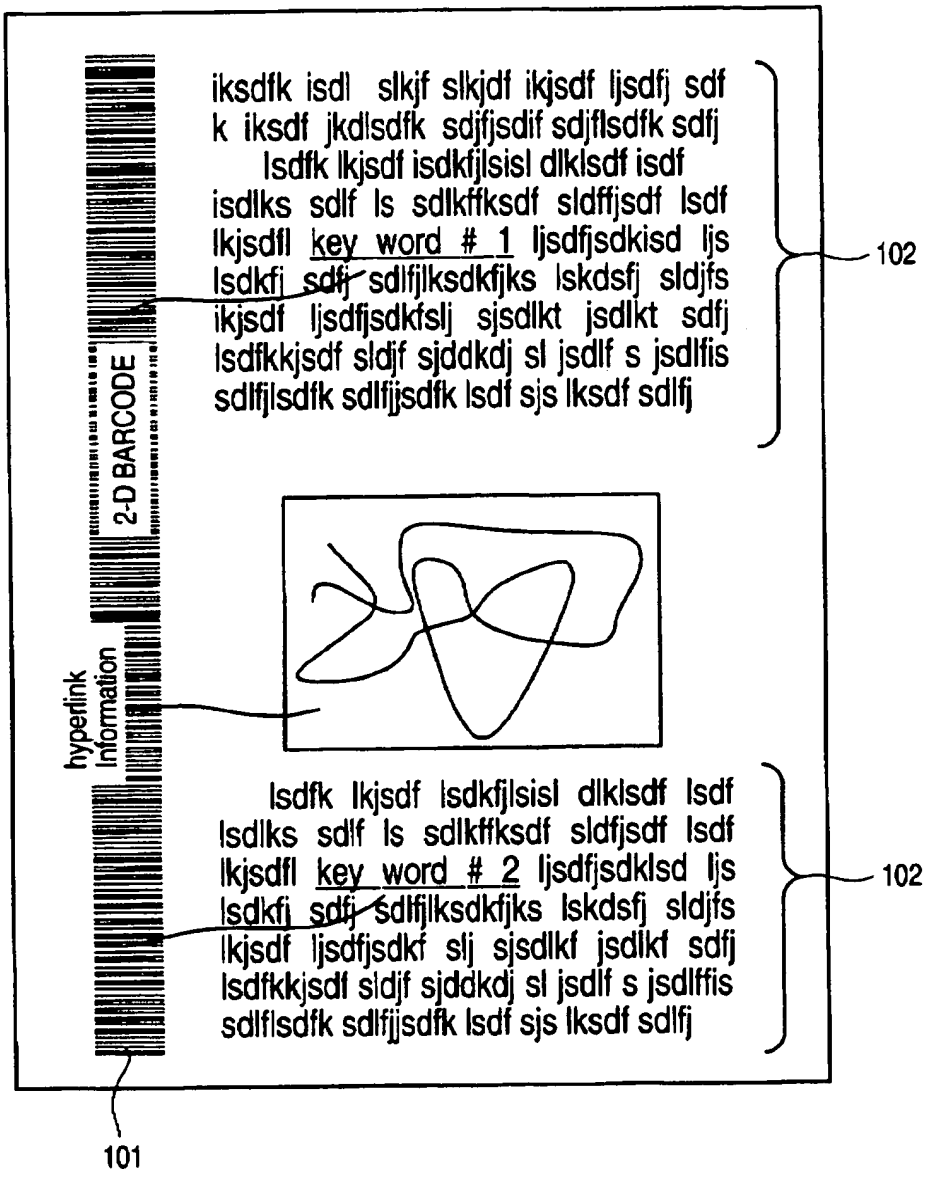

iksdfk isdl slkjf slkjdf ikjsdf ljsdfj sdf
k iksdf jkdlsdfk sdjfjsdif sdjflsdfk sdfj
lsdfk lkjsdf isdkfjlsisl dlklsdf isdf
isdlks sdlf ls sdlkffksdf sldffjsdf lsdf
lkjsdfl key word # 1 ljsdfjsdkisd ljs
lsdkfj sdfj sdlfjlksdkfjks lskdsfj sldjfs
ikjsdf ljsdfjsdkfslj sjsdlkt jsdlkt sdfj
lsdfkkjsdf sldjf sjddkdj sl jsdlf s jsdlfis
sdlfjlsdfk sdlfjjsdfk lsdf sjs lksdf sdlfj
102
2-D BARCODE
hyperlink information
lsdfk lkjsdf lsdkfjlsisl dlklsdf lsdf
lsdlks sdlf ls sdlkffksdf sldfjsdf lsdf
lkjsdfl key word # 2 ljsdfjsdklsd ljs
lsdkfj sdfj sdlfjlksdkfjks lskdsfj sldjfs
lkjsdf ljsdfjsdkf slj sjsdlkf jsdlkf sdfj
lsdfkkjsdf sldjf sjddkdj sl jsdlf s jsdlffis
sdlflsdfk sdlfjjsdfk lsdf sjs lksdf sdlfj
102
101

FIG_1

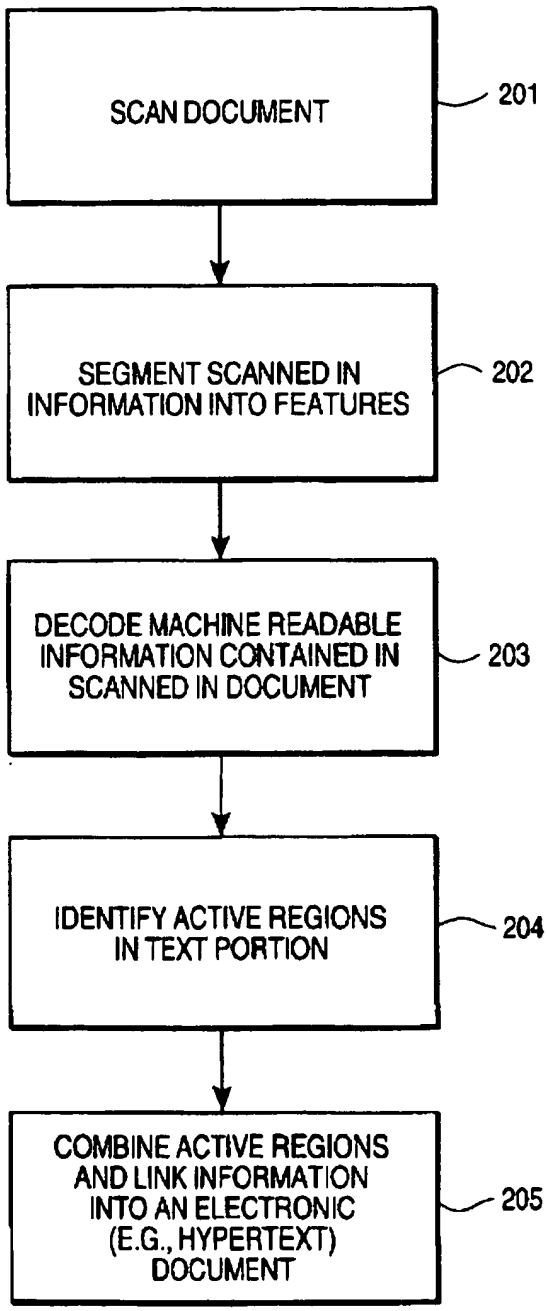
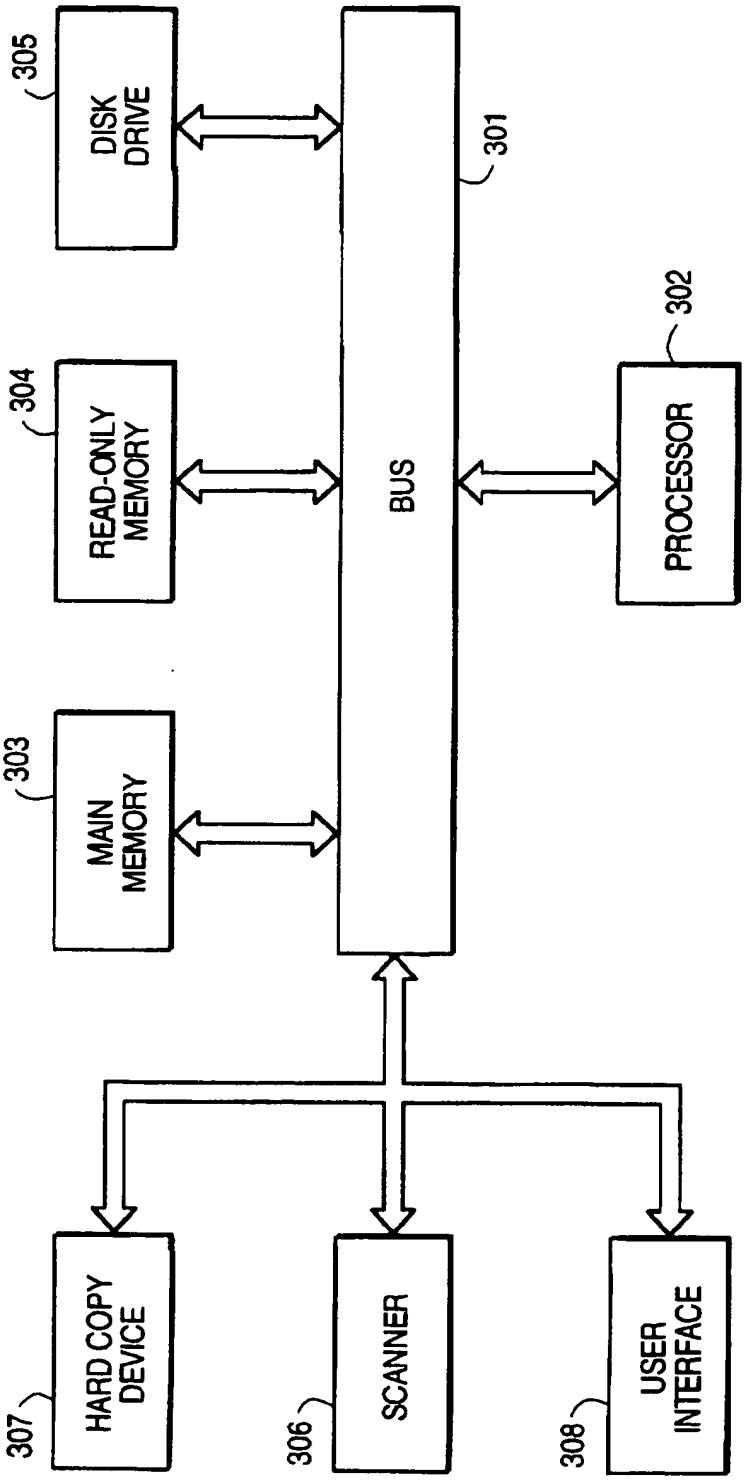

SCAN DOCUMENT
201
SEGMENT SCANNED IN INFORMATION INTO FEATURES
202
DECODE MACHINE READABLE INFORMATION CONTAINED IN SCANNED IN DOCUMENT
203
IDENTIFY ACTIVE REGIONS IN TEXT PORTION
204
COMBINE ACTIVE REGIONS AND LINK INFORMATION INTO AN ELECTRONIC (E.G., HYPERTEXT) DOCUMENT
205

FIG_2


305
DISK DRIVE
301
BUS
302
PROCESSOR
304
READ-ONLY MEMORY
303
MAIN MEMORY
307
HARD COPY DEVICE
306
SCANNER
308
USER INTERFACE

FIG—3

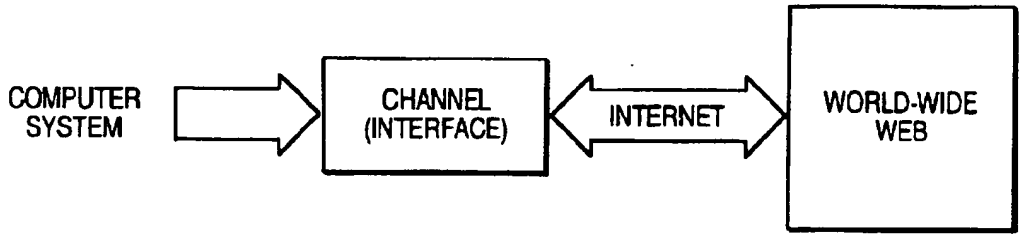

COMPUTER SYSTEM
CHANNEL (INTERFACE)
INTERNET
WORLD-WIDE WEB

FIG_4

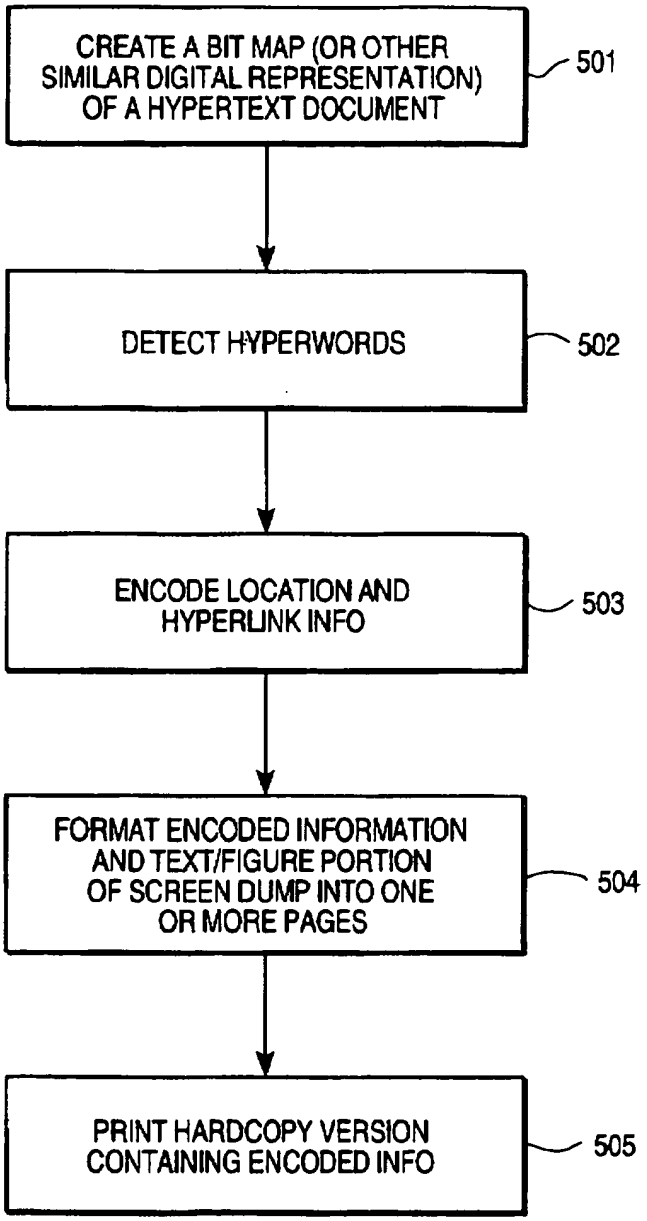

CREATE A BIT MAP (OR OTHER SIMILAR DIGITAL REPRESENTATION) OF A HYPERTEXT DOCUMENT
501
DETECT HYPERWORDS
502
ENCODE LOCATION AND HYPERLINK INFO
503
FORMAT ENCODED INFORMATION AND TEXT/FIGURE PORTION OF SCREEN DUMP INTO ONE OR MORE PAGES
504
PRINT HARDCOPY VERSION CONTAINING ENCODED INFO
505

FIG_5

The RICOH California Research Center

The Art of Imagination

The Science of Information

(This document is still under construction. Links may be broken at the moment. CRC members please try using our Internal server)

Welcome to CRC

At the California Research Center, we are concerned with the science of information. We work throughout the R&D spectrum, from basic science througfh applied research to advanced prototyping.

The California Research Center was formally established in 1989. Our history, charter, location and other background information can be found by clicking here.

CRC Publications

FIG_6A

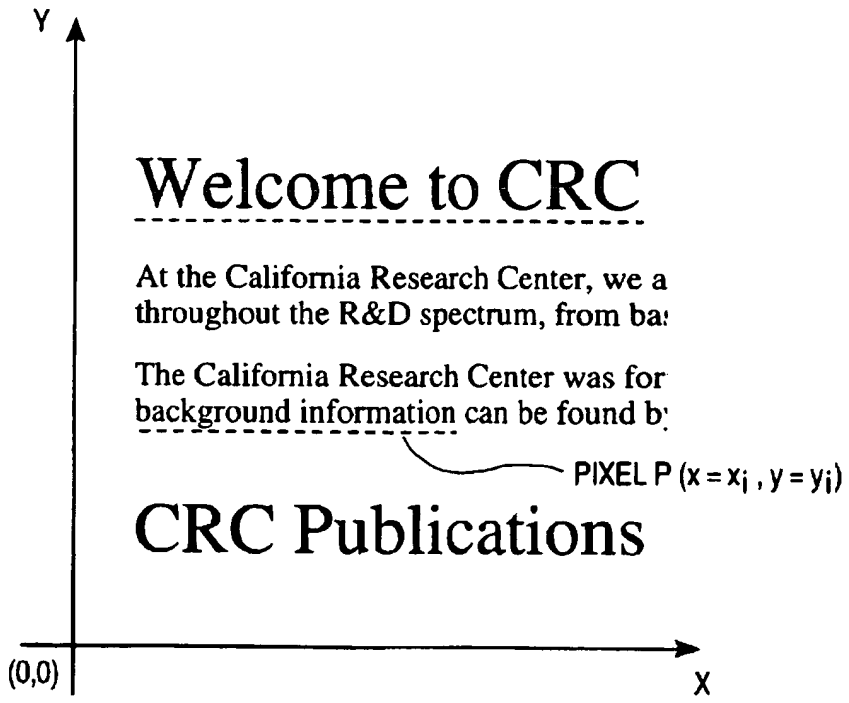

Y
Welcome to CRC
At the California Research Center, we a
throughout the R&D spectrum, from ba
The California Research Center was for
background information can be found b
PIXEL P (x = xi , y = yi)
CRC Publications
(0,0)
X

FIG_6B

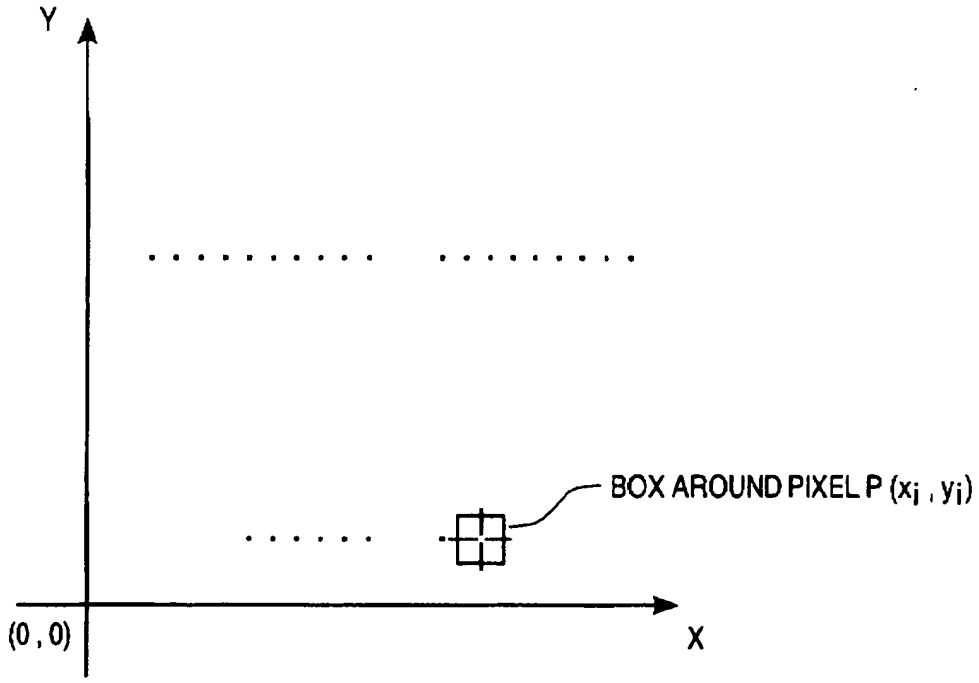


FIG_6C

FIG_6D

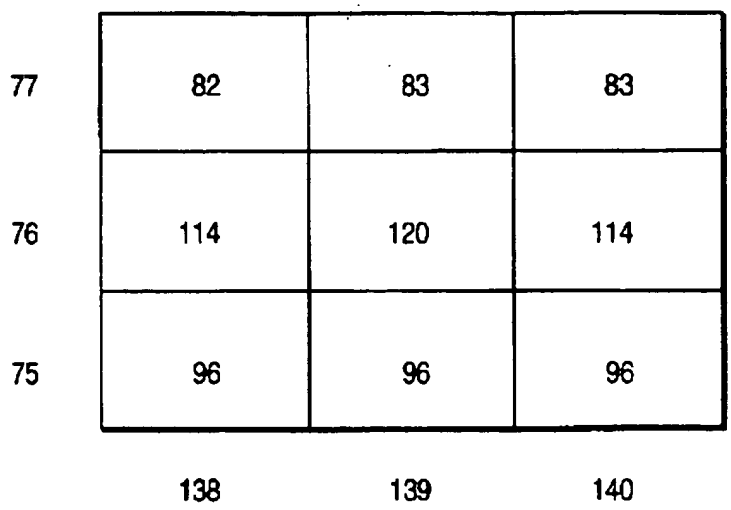


FIG_6E

CONVERTING DOCUMENTS, WITH LINKS TO OTHER ELECTRONIC INFORMATION, BETWEEN HARDCOPY AND ELECTRONIC FORMATS

FIELD OF THE INVENTION

The present invention relates to the field of document conversion and publication; more particularly, the present invention relates to hardcopy document to hyperdocument conversion that permits the construction of a hyperdocument complete with links from a hardcopy version of the document, and vice versa.

BACKGROUND OF THE INVENTION

The "World Wide Web" (hereinafter referred to as "the Web") is a term that describes the interconnected, on-line documents that can be accessed via computer systems hooked to the Internet using software clients. In the prior art, these software clients are graphical browsers, such as Mosaic and Netscape, that permit a user to select various documents. Upon selection, a graphical browser retrieves the documents and provides them to the user, either by displaying them on a display screen or by causing them to be printed on a hardcopy device, such as a printer, or in case where the linked document is an audio file or a movie file, the browser enables appropriate vendition.

Portions of documents displayed using the graphical browser contain hypertext links. The hypertext links link graphics or text on one document with another document on the Web. Documents containing hypertext links are created prior to their "publishing" on the Web. That is, a document that is to be published is provided to a server which creates the document and, essentially, publishes the document by permitting its access by others on the Web. Each hypertext link is associated with a Uniform Resource Locator (URL) that identifies and locates a document on the Web. When a user selects a hypertext link, using, for instance, a cursor, the graphical browser retrieves the corresponding (or linked) document(s).

In order to create documents containing a hypertext link, an individual must specify the specific portion, or active region, of the document that will be responsive to selection by the user. An individual also must identify the specific hypertext link information, including the URL, to set forth the resource or document that is to be located on the Web when the particular active region is selected by the user. Once both the active portions of the document and their corresponding hypertext links have been selected, an application, such as Mosaic, processes the information and produces the document. Note that such a process is not automatic and requires user interaction in order to specify these specific locations and hypertext links. It is desirable to automate such a process to allow the user to create a document having hypertext links without interactively specifying both active regions and their links. That is, it is desirable to have a passive hyperdocument creation system in which hypertext documents may be created automatically.

If a user desires to print a hypertext document, a hardcopy of the displayed document may be printed. However, the printed version does not contain the hypertext information. That is, the link information associated with the document is lost. Furthermore, because the link information is lost, it is impossible to use the printed copy of the document to recreate the hypertext document. If one desires to recreate the hypertext document, the user must again specify both active regions and their link information. It is desirable to automate the process to allow a hypertext document to be printed out and then, using that printout, recreate a hypertext document without having to actively interact in the set up of the hypertext document.

The present invention provides for conversion between hardcopy (e.g., plain text) and hypertext documents. In the present invention, this conversion is performed automatically. Furthermore, the present invention allows a hardcopy version of the hypertext document to be created, such that the link information remains with the hardcopy version. In this manner, the hardcopy document may be used to recreate a hypertext version of the document with the correct, or same, link information.

SUMMARY OF THE INVENTION

The present invention provides a method and apparatus for generating electronic documents from a hardcopy document, and vice versa. A hardcopy document contains encoded link information and one or more regions designated to be active that are associated with the encoding link information. The hardcopy document is scanned and the scanned information is converted into an electronic version of the hardcopy document having active regions. Each active region is linked to electronic information, such that selection of an active region accesses linked electronic information.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of various embodiments of the invention, which, however, should not be taken to limit the invention to the specific embodiments, but are for explanation and understanding only.

FIG. 1 illustrates an example of a hardcopy document containing encoded link information.

FIG. 2 is a flow diagram of one embodiment of the hardcopy to hypertext document conversion process of the present invention.

FIG. 3 is a block diagram of an exemplary computer system that may be employed by the present invention to convert hardcopy versions of documents to electronic versions of documents and vice versa.

FIG. 4 is one embodiment of World Wide Web document retrieval and publishing system of the present invention.

FIG. 5 is a flow diagram of one embodiment of the hypertext version to hardcopy version conversion process of the present invention.

FIG. 6A is an image of a sample hypertext document.

FIG. 6B is a magnified portion of the hypertext represented in FIG. 6A.

FIG. 6C illustrates the hyperwords being detected through the use of a template.

FIG. 6D is a magnified view of one embodiment of a template for locating the hyperwords in a hypertext document.

FIG. 6E illustrates a set of correlation values that result from applying the template of FIG. 6D throughout the hypertext document at the boxed region as shown in FIG. 6C.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

A method and apparatus for converting between a hypertext and plain text document is described. In the following

detailed description of the present invention numerous specific details are set forth, such as number of pages of documents, types of encoding, numbers of links, etc., in order to provide a thorough understanding of the present invention. However, it will be apparent to one skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

Some portions of the detailed descriptions which follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussions, it is appreciated that throughout the present invention, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

The present invention also relates to apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose machines may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these machines will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

Overview of the Present Invention

The present invention provides for converting a hardcopy version of a document containing, for instance, text and/or graphical features such as figures, into an electronic version of the document. In one embodiment, a paper document is converted into an electronic document and vice versa.

One application of the present invention is to convert paper documents into hypertext document. The paper document includes hypertext words and graphical features as well as encoded link information. By decoding the encoded link information and determining the location of active or "clickable" regions in the document, an electronic version of the document may be created.

For example, in the present invention, the user may take a hardcopy document containing, for example, ten pages of multi-column text, tables, and graphics containing "highlighted words", in-text references and a sidechannel of encoded information. The "highlighted" portions symbolize active regions that are designated to be active in the document. The user scans these pages into a document conversion system to form a full electronic hypertext version of the document automatically, having "clickable" links within the document that may link to other portions of the document as well as to a data base of documents, possibly via Mosaic.

The present invention also provides for printing a full electronic (e.g., hypertext) version of a document that is in a system into one or more pages of a hardcopy document such as paper, film, or other similar medium containing text, tables, highlighted words and graphical features, in-text references and a sidechannel of encoded information. This printed document then may be inserted into the document conversion system of the present invention to recreate the electronic (e.g., hypertext) version of the document. Thus, an electronic version of a document on a computer system may be printed to create a hardcopy version, complete with encoded information to enable the recreation of the electronic version.

Plain Text to Hypertext Conversion Process

The plain text to hypertext conversion process of the present invention is performed on a hardcopy (e.g., paper document). An exemplary paper page of the document is shown in FIG. 1. Referring to FIG. 1, the page is shown with text **102** on the right side containing text, key words, and graphical features such as figures. A section of encoded data **101** on the hardcopy document contains information associated with the key words, figures, etc. and their hyperlinks. In other words, the encoded information **101** contains hyperlink information that is used in the document to actively retrieve other portions of the document or other documents.

In one embodiment, the encoded information **101** comprises machine readable information. The machine readable information may be a digital format, such as a barcode or digital paper, one embodiment of which is described in U.S. Pat. No. 5,337,362, entitled "Method and Apparatus for Placing Data Onto Plain Paper", issued on Aug. 9, 1994, and assigned to the corporate assignee of the present invention. Other types of encoding may be used in the present invention as long as identification and selection of links and their association with the "hypertext" (or active regions) in the text are possible.

In one embodiment, the encoded information may be encrypted using one of the well-known encryption techniques, such as public/private key, RSA, etc.

The encoded machine-readable information **101** may be located anywhere on the document where human-readable information is not obscured (This location is referred to as a side channel). In one embodiment, the encoded information **101** is located at the left-handed side of each page. Note that in this case, it is the responsibility of the recognition hardware and/or software in a computer system to locate and identify the machine readable information in order to obtain the hyperlink information contained therein.

As shown in FIG. 1, the encoded information is contained in a boxed region. The encoded information **101** may be contained in one or more (e.g., 2, 3, 4, etc.) boxes.

The information contained in the encoded information **101** is associated with key words and figures and their hyperlinks. In one embodiment, the encoded information **101** for text may identify the location of the key words in the text portion and its associated link information specifying the document (or portion thereof) that is to be retrieved upon selection of the active region. In one embodiment, the key word is actually included in the encoded information. In such case it would be the responsibility of the conversion system to locate that key word in the text, via search, to associate its hyperlink information. In another embodiment, the encoded information **101** may not contain the key word, but instead contain its location. Its location may be specified by the location of the start of a key word or may be in the form of a bounding box which identifies an area in the text portion containing the key word associated with a particular hyperlink. As for figures, the information for a particular figure may only need to encode the rough position of the figure, since a segmentation algorithm, as described later, will locate its true boundaries. Likewise, the code used for a particular word may not specify the precise (x,y) position of the target word. Instead, encoding merely the line number or the order of the key word may suffice while another algorithm is used to identify the correct word. Note, the key word may be located based on the manner in which it is highlighted (e.g., underlining, bold, different color, etc.) in the hardcopy document. In another embodiment, the encoded data might contain links (the actual path information) to audio or movie files, if not just text or graphics.

FIG. **2** illustrates the process of converting a hardcopy document, such as a plain text document, to an electronic document such as a hypertext document. The process of plain text to hypertext document conversion begins by scanning a document having one or more pages (processing block **201**). In one embodiment, each of the pages may resemble the page shown in FIG. **1**. A scanner may be used to perform the function with the resulting output being, for example, a bit map image.

The bit map image is segmented to separate and locate features such as figures, paragraphs, lines, words, letters, and tables from the encoded information (processing block **202**). In one embodiment, the segmentation algorithm may divide the bit map image based on knowledge of where the encoded information is on the page. For instance, a segmentation algorithm may assume that the encoded information is always located on the left side of each page at a specific location.

In one embodiment, the sidechannel information is stored in a box or rectangular shaped area wherein the border may be a certain number of pixels thick. In such a case, the segmentation algorithm performs a search for a border by scanning a row until a predetermined number of black pixels are found in a row. In one embodiment, a border may be three pixels thick. Once the box has been located, the information contained therein may be decoded. This decoding may be performed in a manner similar to that described in U.S. Pat. No. 5,337,362, entitled "Method and Apparatus for Placing Data Onto Plain Paper", issued Aug. 9, 1994, and assigned to the corporate assignee of the present invention.

Once the location of the encoded link information has been identified, the machine readable information is recognized and decoded to determine the hypertext link information contained therein (processing block **203**).

Next, the "highlighted" regions (i.e., regions designated to be active) are identified (processing block **204**). In one embodiment, the encoded information indicates the location of each active region. That is, active words or other elements may be identified by specific information setting forth the location and size (e.g., specifying a bounding box within which the word or object exists) in the machine readable information in the sidechannel. In an alternate embodiment, the identification of these active regions is made using one or more templates. The template may represent the form of highlighting used to set off the word or figure in the plain text. The template may be a bitmapped version of the highlighting. For instance, alternating black and white dashes may be used to underline each of the key words. In the present invention, an algorithm uses a bit mapped version of the alternating black and white dashes as a template to search the plain text until a strong correlation or perfect match occurs between the template and information in the text portion. If there is a correlation between the template and the area being checked that is above a specified threshold, then it is determined that hypertext link or word exists. The template(s) may be specified in the encoded sidechannel information or may be selected by the recognition algorithm. Note that the use of thresholds is well-known in the art. Further, the thresholds may be set by a system designer and/or user.

It should be noted that recognition of the digital information may be performed after identifying specific graphical features, paragraphs, lines, words, letters, tables in the text portion of the scanned hardcopy document.

Once the encoded information has been decoded and specific active regions in the plain text portion of the document identified, the hyperlink information is associated with the active words and/or objects (processing block **205**), and processing is performed to create the hypertext document in a manner well-known in the art, such that the selection of one of the marked words causes that portion of the document or other documents to be retrieved based on a resource locator within the link information.

Note that when creating the hypertext document image to be displayable on screen, the active regions (e.g., the key words or objects) associated with hyperlink information may be highlighted in any manner. The manner of highlighting may be specified by the processing procedure creating the hypertext document. The highlighting may be underlining or placing the active portions of a document in a different font or color (e.g., bold). It should be noted that the highlighting selected for the newly created hypertext document does not have to be the same form of highlighting used to designate the active regions in the hardcopy version of the document.

FIG. **3** illustrates an overview of one embodiment of the document conversion system of the present invention shown in block diagram form. The system of the present invention is a digital processing system. It will be understood that while FIG. **3** is useful for providing an overall description of the processing system of the present invention, a number of details of the system are not shown. As necessary for disclosure of the present invention, further detail is set forth with other figures provided with this specification. Further, the present invention is described with reference to various embodiment; alternative embodiments which may be conceived by one of ordinary skill in the art are considered within the scope of the claims set forth below.

Referring to FIG. **3**, the system of the present invention includes a bus or other communication means **301** for communicating information. Processor **302** is coupled with bus **301** for processing information. A random access memory (RAM) or other dynamic storage device **303** (commonly referred to as a main memory) for storing information and instructions for processor **302** is also coupled to bus **301**. Also coupled to bus **302** is a read only memory (ROM) or other static storage device **304** for storing static information and instructions for processor **302** and a data storage device **305**, such as a magnetic disk and disk drive for storing information and instructions.

The processing system also includes a scanner **306** coupled to bus **301** for scanning selected hardcopy versions of documents into the system. Scanner **306** is capable of reading digital representations of images (e.g., digital paper), as well as regular images.

In one embodiment, scanner **306** comprises a gray scale scanner. In one embodiment, the resolution of scanner **306** resolution is 200 DPI. Scanner **306** converts the individual picture elements, referred to as pixels, of the scanned image into digital values. In one embodiment, scanner **306** comprises an Image Scanner ICS-400 brand scanner of Ricoh Corporation of West Caldwell, N.J. In other embodiments, scanner **306** is a bit map scanner which scans the image of each hardcopy input document in a predetermined spatial resolution to produce digital values. These digital values collectively produce a data structure known as a bit map image, which is well-known to those in the art.

Note that in one embodiment, processor **302** runs one or more software routines to perform the segmentation, decoding and electronic document creation processes for a document conversion system using the input from scanner **306**. Thus, in one embodiment, processor **302** operates as the encoder and the decoder of the present invention. The software may be stored and accessed from a memory in the system. Note also that processor **302** may also perform encryption and decryption if employed in the document conversion process.

The computer system may include separate decoder hardware and encoder hardware distinct from processor **302**.

A hardcopy device **307** is also coupled to bus **301** for printing hard copy versions of documents that were converted from electronic formats. In one embodiment, hardcopy device **307** could comprise a plotter or printer, such as a bit map printer that maps the digital values of a bit map image into pixels which are printed on plain paper, film, or other similar media.

Moreover, a human or user interface **308** is included for enabling a user to interact with processor **302**, scanner **306** and hardcopy device **307**. User interface **308** represents the input and output devices through which the user enters control instructions to and receives feedback from the photocopier (i.e., processor **302**, scanner **306** and hardcopy device **307**).

User interface **308** could include an alphanumeric input device including alphanumeric and other keys for communicating information and command selections, a cursor control device for controlling cursor movement and/or a display device, such as a cathode ray tube (CRT), liquid crystal display, etc. for displaying information to the photocopier user. Note that these components are well-known in the art and have been omitted to avoid unnecessarily obscuring the present invention. Note that in one embodiment, user interface **308** may provide inputs directly to scanner **306** and hardcopy device **307** which can be distinct from those inputs to processor **302**.

Note that these components may be integrated together into a single system or separately accessible components, such as a scanner **306** and hardcopy device **307**. In one embodiment, scanner **306** and/or hardcopy device **307** could be coupled to bus **301** using dedicated communication links or switchable communication networks.

In one embodiment, the processor **302**, in conjunction with the rest of the computer system, runs conversion software that performs the following functions. First, the conversion software decodes the machine readable information that is encoded in the sidechannel to obtain link information to associate the links with highlighted, or active, words or objects in the text portion. Second, the conversion software recognizes words or active regions in the text portion. This further could be performed using location information specified in the machine readable information (encoded portion) of the bit map image of the plain paper document or using a template specified in the decoded machine readable information or defined in storage and always used by the conversion software. Third, the conversion software translates the link information and the recognized highlighted regions into a hypertext document. Note that the performance of each of these conversion step is well-known to those skilled in the art.

In one embodiment, the computer system is coupled via channel or other data transfer mechanism to a Web site in the World Wide Web (WWW). Such a connection is illustrated in FIG. **4**. In one embodiment, the computer system may comprise the Web site and the Web site is coupled to the Internet or other network or document resource which provides access to the WWW.

Once a hypertext document has been created, it may be published for use on the WWW with the Internet, such that selection of active regions (hyperwords) on an electronic version of the document being displayed causes the retrieval of other documents or portions of the document that may be located by using the universal resource locator or other resource identifier.

It should be noted that the present invention is not limited to retrieving and publishing documents on the WWW or the Internet. The teachings of the present invention may be applied to various networks, data document storage and archival facilities, or other types of client/service systems which have documents or other information available upon request.

The present invention includes a process by which a hypertext document is converted into a plain paper document. One embodiment of this process is shown in FIG. **5**. The hardcopy document that results contains hypertext link information in machine readable format to enable conversion back into a hypertext document format. Thus, the link information will be available to the user to enable a reversal back into a hypertext document.

Referring to FIG. **5**, the conversion process begins by creating a bit map of a hyperpage that is currently displayed on the display screen from a screen "dump" (processing block **501**). An example of such a document in shown in FIG. **6A**. A portion of the bit map is shown in FIG. **6B**.

Once a bit map has been created, the present invention detects the hyperwords in the bit map (processing block **502**). In one embodiment, the hyperwords are detected by using a template. Such a template is shown in FIG. **6D**. The present invention searches the bit map locating portions of the bit map with the template, generating correlation values.

FIG. **6B** illustrates correlation values around the boxed region shown in FIG. **6C**. The correlation value indicate whether there is a high correlation between portions of the document and the template.

Once the marked words have been identified, the location and hyperlink information is encoded (processing block **503**), the information is formatted into one or more pages (processing block **504**), and a hardcopy document is printed having a sidechannel with encoded hyperlink information (processing block **505**). In one embodiment, a template may be encoded as well for use in locating active regions.

Whereas, many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after having read the foregoing description, it is to be understood that the particular embodiment shown and described by way of illustration are in no way to be considered limiting. Therefore, reference to the details of the various embodiments are not intended to limit the scope of the claims which themselves recite only those features regarded as essential to the invention.

Thus, a method and apparatus for converting between hardcopy versions and electronic versions of documents has been described.

We claim:

1. A method of converting a hardcopy document to an electronic document comprising the steps of:

scanning a hardcopy document containing encoded link information and one or more regions designated to be active associated with said encoded link information;

decoding the encoded link information to obtain the link information associated with said one or more regions;

locating said one or more regions in a scanned-in version of the hardcopy document, wherein the step of locating comprises searching a bit map of the scanned hardcopy document using a template stored in the encoded link information and obtained when decoding the encoded link information; and

creating an electronic version of the hardcopy document having said one or more regions linked to electronic information, such that selection of any of said one or more regions accesses linked electronic information.

2. The method defined in claim 1 wherein the encoded link information identifies a location of at least one of said one or more regions in the document.

3. The method defined in claim 1 wherein the encoded link information includes location information identifying a location of said at least one of said one or more regions in the document.

4. A method of generating an electronic document comprising the steps of:

scanning a hardcopy document containing encoded link information and one or more regions designated to be active associated with said encoded link information; and

converting scanned information to an electronic version of the hardcopy document having active regions, wherein the step of converting comprises

decoding the encoded link information,

determining a location of each of said active regions, and

linking each active region to electronic information via decoded link information, such that selection of an active region accesses linked electronic information, and

wherein the encoded link information includes location information identifying a location of at least one of said active regions in the document, and further wherein the location information identifies at least one word located in a text portion of the hardcopy document associated with said at least one active region.

5. The method defined in claim 4 wherein the location information identifies a start of a word associated with said at least one active region.

6. The method defined in claim 4 wherein the scanned information is converted a hypertext version of the paper document having hypertext links one word located in a text portion of the hardcopy document associated with.

7. The method defined in claim 6 wherein the location information identifies a start of a word associated with said at least one active region.

8. A method of generating an electronic document comprising the steps of:

scanning a hardcopy document containing encoded link information and one or more regions designated to be active associated with said encoded link information; and

converting scanned information to an electronic version of the hardcopy document having active regions, wherein the step of converting comprises

decoding the encoded link information,

determining a location of each of said active regions, and

linking each active region to electronic information via decoded link information, such that selection of an active region accesses linked electronic information, and

wherein the encoded link information includes location information identifying a location of at least one of said active regions in the document, and further wherein the location information indicates a bounding box specifying an area containing said at least one active region.

9. The method defined in claim 8 wherein the scanned information is converted a hypertext version of the paper document having hypertext links.

10. A method of generating an electronic document comprising the steps of:

scanning a hardcopy document containing encoded link information and one or more regions designated to be active associated with said encoded link information; and

converting scanned information to an electronic version of the hardcopy document having active regions, wherein the step of converting comprises

decoding the encoded link information,

determining a location of each of said active regions, and

linking each active region to electronic information via decoded link information, such that selection of an active region accesses linked

wherein the encoded link information includes location information identifying a location of at least one of said active regions in the document, and further wherein the location information indicates a line number and order location associated with said at least one active region.

11. The method defined in claim 8 wherein the scanned information is converted a hypertext version of the paper document having hypertext links.

12. A method of generating an electronic document comprising the steps of:

scanning a hardcopy document containing encoded link information and one or more regions designated to be active associated with said encoded link information; and

converting scanned information to an electronic version of the hardcopy document having active regions, wherein the step of converting comprises

decoding the encoded link information,

determining a location of each of said active regions, and

linking each active region to electronic information via decoded link information, such that selection of an active region accesses linked electronic information, and

wherein the encoded link information sets forth at least one of said active regions located in the document and the manner in which said at least one active region is highlighted.

13. The method defined in claim 12 wherein the encoded link information identifies a location of said at least one of said active regions in the document.

14. The method defined in claim 12 wherein the encoded link information includes location information identifying a location of said at least one of said active regions in the document.

15. The method defined in claim 12 wherein the scanned information is converted a hypertext version of the paper document having hypertext links.

16. A method of generating an electronic document comprising the steps of:

scanning a hardcopy document containing encoded link information and one or more regions designated to be active associated with said encoded link information; and

converting scanned information to an electronic version of the hardcopy document having active regions, wherein the step of converting comprises

decoding the encoded link information,

determining a location of each of said active regions, wherein the step of determining a location of each of said active regions comprises searching a bit map of the scanned hardcopy document using a template that is stored in the encoded link information and obtained when decoding the encoded link information, and

linking each active region to electronic information via decoded link information, such that selection of an active region accesses linked electronic information.

17. The method defined in claim 16 wherein the scanned information is converted a hypertext version of the paper document having hypertext links.

18. A system for converting a hardcopy version of a document to an electronic version comprising:

a scanner to scan the hardcopy version of the document;

a decoder coupled to the scanner to decode encoded link information in scanned information from the scanner, wherein the encoded link information includes location information identifying a location of at least one active region in the document, and further wherein the location information identifies at least one word located in a text portion of the hardcopy document associated with said at least one active region; and

processing logic coupled to the scanner and the decoder to identify and locate active regions designated to be linked to other electronic information indicated in the encoded link information and to create the electronic version in which selection of one of said active regions when said electronic version is displayed causes the linked information associated with said one of said active regions to be retrieved.

19. The apparatus defined in claim 18 wherein the location information identifies a start of a word associated with said at least one active region.

20. A system for converting a hardcopy version of a document to an electronic version comprising:

a scanner to scan the hardcopy version of the document;

a decoder coupled to the scanner to decode encoded link information in scanned information from the scanner, wherein the encoded link information includes location information identifying a location of at least one active region in the document, and further wherein the location information indicates a bounding box specifying an area containing said at least one active region; and

processing logic coupled to the scanner and the decoder to identify and locate active regions designated to be linked to other electronic information indicated in the encoded link information and to create the electronic version in which selection of one of said active regions when said electronic version is displayed causes the linked information associated with said one of said active regions to be retrieved.

21. A system for converting a hardcopy version of a document to an electronic version comprising:

a scanner to scan the hardcopy version of the document;

a decoder coupled to the scanner to decode encoded link information in scanned information from the scanner, wherein the encoded link information includes location information identifying a location of at least one active region in the document, and further wherein the location information indicates a line number and order location associated with said at least one active region; and

processing logic coupled to the scanner and the decoder to identify and locate active regions designated to be linked to other electronic information indicated in the encoded link information and to create the electronic version in which selection of one of said active regions when said electronic version is displayed causes the linked information associated with said one of said active regions to be retrieved.

22. A system for converting a hardcopy version of a document to an electronic version comprising:

a scanner to scan the hardcopy version of the document;

a decoder coupled to the scanner to decode encoded link information in scanned information from the scanner, wherein the encoded link information sets forth at least one of said active regions located in the document and the manner in which said at least one active region is highlighted; and

processing logic coupled to the scanner and the decoder to identify and locate active regions designated to be linked to other electronic information indicated in the encoded link information and to create the electronic version in which selection of one of said active regions when said electronic version is displayed causes the linked information associated with said one of said active regions to be retrieved.

23. A system for converting a hardcopy version of a document to an electronic version comprising:

a scanner to scan the hardcopy version of the document;

a decoder coupled to the scanner to decode encoded link information in scanned information from the scanner; and

processing logic coupled to the scanner and the decoder to identify and locate regions designated to be linked to other electronic information indicated in the encoded link information and to create the electronic version in which selection of one of said regions when said electronic version is displayed causes the linked information associated with said one of said regions to be retrieved, and wherein the processing logic searches a bit map of the scanned hardcopy document using a template that is stored in the encoded link information and obtained when decoding the encoded link information.

* * * * *